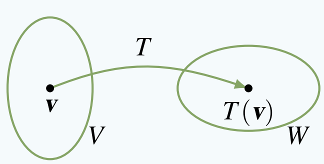
7. Linear Transformations

If V and W are vector spaces, a function $T : V \to W$ is a rule that assigns to each vector $\mathbf{v}$ in V a uniquely determined vector $T(\mathbf{v})$ in W. As mentioned in Section 2.2, two functions $S : V \to W$ and $T : V \to W$ are equal if $S(\mathbf{v}) = T(\mathbf{v})$ for every $\mathbf{v}$ in V. A function $T : V \to W$ is called a *linear transformation* if $T(\mathbf{v}+\mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1)$ for all $\mathbf{v}$, $\mathbf{v}_1$ in V and $T(r\mathbf{v}) = rT(\mathbf{v})$ for all $\mathbf{v}$ in V and all scalars r. $T(\mathbf{v})$ is called the *image* of $\mathbf{v}$ under T. We have already studied linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ and shown (in Section 2.6) that they are all given by multiplication by a uniquely determined $m \times n$ matrix A; that is $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. In the case of linear operators $\mathbb{R}^2 \to \mathbb{R}^2$, this yields an important way to describe geometric functions such as rotations about the origin and reflections in a line through the origin.

In the present chapter we will describe linear transformations in general, introduce the *kernel* and *image* of a linear transformation, and prove a useful result (called the *dimension theorem*) that relates the dimensions of the kernel and image, and unifies and extends several earlier results. Finally we study the notion of *isomorphic* vector spaces, that is, spaces that are identical except for notation, and relate this to composition of transformations that was introduced in Section 2.3.

7.1 Examples and Elementary Properties

Definition 7.1 Linear Transformations of Vector Spaces



If V and W are two vector spaces, a function $T : V \to W$ is called a ***linear transformation*** *if it satisfies the following axioms.*

T1.	$T(\mathbf{v}+\mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1)$	*for all $\mathbf{v}$ and $\mathbf{v}_1$ in V.*
T2.	$T(r\mathbf{v}) = rT(\mathbf{v})$	*for all $\mathbf{v}$ in V and r in $\mathbb{R}$.*

A linear transformation $T : V \to V$ is called a ***linear operator*** *on V. The situation can be visualized as in the diagram.*

Axiom T1 is just the requirement that T *preserves* vector addition. It asserts that the result $T(\mathbf{v}+\mathbf{v}_1)$ of adding $\mathbf{v}$ and $\mathbf{v}_1$ first and then applying T is the same as applying T first to get $T(\mathbf{v})$ and $T(\mathbf{v}_1)$ and then adding. Similarly, axiom T2 means that T *preserves* scalar multiplication. Note that, even though the additions in axiom T1 are both denoted by the same symbol $+$, the addition on the left forming $\mathbf{v}+\mathbf{v}_1$ is carried out in V, whereas the addition $T(\mathbf{v}) + T(\mathbf{v}_1)$ is done in W. Similarly, the scalar multiplications $r\mathbf{v}$ and $rT(\mathbf{v})$ in axiom T2 refer to the spaces V and W, respectively.

We have already seen many examples of linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$. In fact, writing vectors in $\mathbb{R}^n$ as columns, Theorem 2.6.2 shows that, for each such T, there is an $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^n$. Moreover, the matrix A is given by $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$. We denote this transformation by $T_A : \mathbb{R}^n \to \mathbb{R}^m$, defined by

$$T_A(\mathbf{x}) = A\mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

Example 7.1.1 lists three important linear transformations that will be referred to later. The verification of axioms T1 and T2 is left to the reader.

Example 7.1.1

If V and W are vector spaces, the following are linear transformations:

Identity operator $V \to V$	$1_V : V \to V$	where $1_V(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v}$ in V
Zero transformation $V \to W$	$0 : V \to W$	where $0(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in V
Scalar operator $V \to V$	$a : V \to V$	where $a(\mathbf{v}) = a\mathbf{v}$ for all $\mathbf{v}$ in V
		(Here a is any real number.)

The symbol 0 will be used to denote the zero transformation from V to W for *any* spaces V and W. It was also used earlier to denote the zero function $[a, b] \to \mathbb{R}$.

The next example gives two important transformations of matrices. Recall that the trace $\operatorname{tr} A$ of an $n \times n$ matrix A is the sum of the entries on the main diagonal.

Example 7.1.2

Show that the transposition and trace are linear transformations. More precisely,

$$\begin{array}{ll} R : \mathbf{M}_{mn} \to \mathbf{M}_{nm} & \text{where } R(A) = A^T \text{ for all } A \text{ in } \mathbf{M}_{mn} \\ S : \mathbf{M}_{mn} \to \mathbb{R} & \text{where } S(A) = \operatorname{tr} A \text{ for all } A \text{ in } \mathbf{M}_{nn} \end{array}$$

are both linear transformations.

Solution. Axioms T1 and T2 for transposition are $(A+B)^T = A^T + B^T$ and $(rA)^T = r(A^T)$, respectively (using Theorem 2.1.2). The verifications for the trace are left to the reader.

Example 7.1.3

If a is a scalar, define $E_a : \mathbf{P}_n \to \mathbb{R}$ by $E_a(p) = p(a)$ for each polynomial p in $\mathbf{P}_n$. Show that E_a is a linear transformation (called **evaluation** at a).

Solution. If p and q are polynomials and r is in $\mathbb{R}$, we use the fact that the sum $p+q$ and scalar product rp are defined as for functions:

$$(p+q)(x) = p(x) + q(x) \quad \text{and} \quad (rp)(x) = rp(x)$$

for all x. Hence, for all p and q in $\mathbf{P}_n$ and all r in $\mathbb{R}$:

$$\begin{aligned} E_a(p+q) &= (p+q)(a) = p(a) + q(a) = E_a(p) + E_a(q), \quad \text{and} \\ E_a(rp) &= (rp)(a) = rp(a) = rE_a(p). \end{aligned}$$

Hence E_a is a linear transformation.

The next example involves some calculus.

Example 7.1.4

Show that the differentiation and integration operations on $\mathbf{P}_n$ are linear transformations. More precisely,

$$D : \mathbf{P}_n \to \mathbf{P}_{n-1} \quad \text{where } D[p(x)] = p'(x) \text{ for all } p(x) \text{ in } \mathbf{P}_n$$
$$I : \mathbf{P}_n \to \mathbf{P}_{n+1} \quad \text{where } I[p(x)] = \int_0^x p(t)dt \text{ for all } p(x) \text{ in } \mathbf{P}_n$$

are linear transformations.

Solution. These restate the following fundamental properties of differentiation and integration.

$$[p(x)+q(x)]' = p'(x)+q'(x) \quad \text{and} \quad [rp(x)]' = (rp)'(x)$$

$$\int_0^x [p(t)+q(t)]dt = \int_0^x p(t)dt + \int_0^x q(t)dt \quad \text{and} \quad \int_0^x rp(t)dt = r\int_0^x p(t)dt$$

The next theorem collects three useful properties of *all* linear transformations. They can be described by saying that, in addition to preserving addition and scalar multiplication (these are the axioms), linear transformations preserve the zero vector, negatives, and linear combinations.

Theorem 7.1.1

Let $T : V \to W$ be a linear transformation.

1. $T(\mathbf{0}) = \mathbf{0}$.
2. $T(-\mathbf{v}) = -T(\mathbf{v})$ *for all* $\mathbf{v}$ *in* V.
3. $T(r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_k\mathbf{v}_k) = r_1T(\mathbf{v}_1) + r_2T(\mathbf{v}_2) + \cdots + r_kT(\mathbf{v}_k)$ *for all* $\mathbf{v}_i$ *in* V *and all* r_i *in* $\mathbb{R}$.

Proof.

1. $T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$ for any $\mathbf{v}$ in V.
2. $T(-\mathbf{v}) = T[(-1)\mathbf{v}] = (-1)T(\mathbf{v}) = -T(\mathbf{v})$ for any $\mathbf{v}$ in V.
3. The proof of Theorem 2.6.1 goes through. □

The ability to use the last part of Theorem 7.1.1 effectively is vital to obtaining the benefits of linear transformations. Example 7.1.5 and Theorem 7.1.2 provide illustrations.

Example 7.1.5

Let $T : V \to W$ be a linear transformation. If $T(\mathbf{v} - 3\mathbf{v}_1) = \mathbf{w}$ and $T(2\mathbf{v} - \mathbf{v}_1) = \mathbf{w}_1$, find $T(\mathbf{v})$ and $T(\mathbf{v}_1)$ in terms of $\mathbf{w}$ and $\mathbf{w}_1$.

Solution. The given relations imply that

$$\begin{aligned} T(\mathbf{v}) - 3T(\mathbf{v}_1) &= \mathbf{w} \\ 2T(\mathbf{v}) - T(\mathbf{v}_1) &= \mathbf{w}_1 \end{aligned}$$

by Theorem 7.1.1. Subtracting twice the first from the second gives $T(\mathbf{v}_1) = \frac{1}{5}(\mathbf{w}_1 - 2\mathbf{w})$. Then substitution gives $T(\mathbf{v}) = \frac{1}{5}(3\mathbf{w}_1 - \mathbf{w})$.

The full effect of property (3) in Theorem 7.1.1 is this: If $T : V \to W$ is a linear transformation and $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)$ are known, then $T(\mathbf{v})$ can be computed for *every* vector $\mathbf{v}$ in $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. In particular, if $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ spans V, then $T(\mathbf{v})$ is determined for all $\mathbf{v}$ in V by the choice of $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)$. The next theorem states this somewhat differently. As for functions in general, two linear transformations $T : V \to W$ and $S : V \to W$ are called **equal** (written $T = S$) if they have the same **action**; that is, if $T(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v}$ in V.

Theorem 7.1.2

Let $T : V \to W$ and $S : V \to W$ be two linear transformations. Suppose that $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. If $T(\mathbf{v}_i) = S(\mathbf{v}_i)$ for each i, then $T = S$.

Proof. If $\mathbf{v}$ is any vector in $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, write $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ where each a_i is in $\mathbb{R}$. Since $T(\mathbf{v}_i) = S(\mathbf{v}_i)$ for each i, Theorem 7.1.1 gives

$$\begin{aligned} T(\mathbf{v}) &= T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) \\ &= a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \cdots + a_nT(\mathbf{v}_n) \\ &= a_1S(\mathbf{v}_1) + a_2S(\mathbf{v}_2) + \cdots + a_nS(\mathbf{v}_n) \\ &= S(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) \\ &= S(\mathbf{v}) \end{aligned}$$

Since $\mathbf{v}$ was arbitrary in V, this shows that $T = S$. □

Example 7.1.6

Let $V = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Let $T : V \to W$ be a linear transformation. If $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_n) = \mathbf{0}$, show that $T = 0$, the zero transformation from V to W.

Solution. The zero transformation $0 : V \to W$ is defined by $0(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in V (Example 7.1.1), so $T(\mathbf{v}_i) = 0(\mathbf{v}_i)$ holds for each i. Hence $T = 0$ by Theorem 7.1.2.

Theorem 7.1.2 can be expressed as follows: If we know what a linear transformation $T : V \to W$ does to each vector in a spanning set for V, then we know what T does to *every* vector in V. If the spanning set is a basis, we can say much more.

Theorem 7.1.3

Let V and W be vector spaces and let $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be a basis of V. Given any vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ in W (they need not be distinct), there exists a unique linear transformation $T : V \to W$ satisfying $T(\mathbf{b}_i) = \mathbf{w}_i$ for each $i = 1, 2, \dots, n$. In fact, the action of T is as follows: Given $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n$ in V, v_i in $\mathbb{R}$, then

$$T(\mathbf{v}) = T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdots + v_n\mathbf{w}_n.$$

Proof. If a transformation T *does* exist with $T(\mathbf{b}_i) = \mathbf{w}_i$ for each i, and if S is any other such transformation, then $T(\mathbf{b}_i) = \mathbf{w}_i = S(\mathbf{b}_i)$ holds for each i, so $S = T$ by Theorem 7.1.2. Hence T is unique if it exists, and it remains to show that there really is such a linear transformation. Given $\mathbf{v}$ in V, we must specify $T(\mathbf{v})$ in W. Because $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V, we have $\mathbf{v} = v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n$, where $v_1, \dots, v_n$ are *uniquely* determined by $\mathbf{v}$ (this is Theorem 6.3.1). Hence we may define $T : V \to W$ by

$$T(\mathbf{v}) = T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdots + v_n\mathbf{w}_n$$

for all $\mathbf{v} = v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n$ in V. This satisfies $T(\mathbf{b}_i) = \mathbf{w}_i$ for each i; the verification that T is linear is left to the reader. □

This theorem shows that linear transformations can be defined almost at will: Simply specify where the basis vectors go, and the rest of the action is dictated by the linearity. Moreover, Theorem 7.1.2 shows that deciding whether two linear transformations are equal comes down to determining whether they have the same effect on the basis vectors. So, given a basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of a vector space V, there is a different linear transformation $V \to W$ for every ordered selection $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ of vectors in W (not necessarily distinct).

Example 7.1.7

Find a linear transformation $T : \mathbf{P}_2 \to \mathbf{M}_{22}$ such that

$$T(1+x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad T(x+x^2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad T(1+x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Solution. The set $\{1+x, x+x^2, 1+x^2\}$ is a basis of $\mathbf{P}_2$, so every vector $p = a + bx + cx^2$ in $\mathbf{P}_2$ is a linear combination of these vectors. In fact

$$p(x) = \tfrac{1}{2}(a+b-c)(1+x) + \tfrac{1}{2}(-a+b+c)(x+x^2) + \tfrac{1}{2}(a-b+c)(1+x^2)$$

Hence Theorem 7.1.3 gives

$$\begin{aligned} T[p(x)] &= \tfrac{1}{2}(a+b-c)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \tfrac{1}{2}(-a+b+c)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \tfrac{1}{2}(a-b+c)\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \tfrac{1}{2}\begin{bmatrix} a+b-c & -a+b+c \\ -a+b+c & a-b+c \end{bmatrix} \end{aligned}$$

Exercises for 7.1

Exercise 7.1.1 Show that each of the following functions is a linear transformation.

a. $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (x, -y)$ (reflection in the x axis)

b. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x, y, -z)$ (reflection in the x-y plane)

c. $T : \mathbb{C} \to \mathbb{C}$; $T(z) = \overline{z}$ (conjugation)

d. $T : \mathbf{M}_{mn} \to \mathbf{M}_{kl}$; $T(A) = PAQ$, P a $k \times m$ matrix, Q an $n \times l$ matrix, both fixed

e. $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$; $T(A) = A^T + A$

f. $T : \mathbf{P}_n \to \mathbb{R}$; $T[p(x)] = p(0)$

g. $T : \mathbf{P}_n \to \mathbb{R}$; $T(r_0 + r_1x + \cdots + r_nx^n) = r_n$

h. $T : \mathbb{R}^n \to \mathbb{R}$; $T(\mathbf{x}) = \mathbf{x} \cdot \mathbf{z}$, $\mathbf{z}$ a fixed vector in $\mathbb{R}^n$

i. $T : \mathbf{P}_n \to \mathbf{P}_n$; $T[p(x)] = p(x+1)$

j. $T : \mathbb{R}^n \to V$; $T(r_1, \cdots, r_n) = r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n$ where $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a fixed basis of V

k. $T : V \to \mathbb{R}$; $T(r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n) = r_1$, where $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a fixed basis of V

Exercise 7.1.2 In each case, show that T is *not* a linear transformation.

a. $T : \mathbf{M}_{nn} \to \mathbb{R}$; $T(A) = \det A$

b. $T : \mathbf{M}_{nm} \to \mathbb{R}$; $T(A) = \operatorname{rank} A$

c. $T : \mathbb{R} \to \mathbb{R}$; $T(x) = x^2$

d. $T : V \to V$; $T(\mathbf{v}) = \mathbf{v} + \mathbf{u}$ where $\mathbf{u} \neq \mathbf{0}$ is a fixed vector in V (T is called the **translation** by $\mathbf{u}$)

Exercise 7.1.3 In each case, assume that T is a linear transformation.

a. If $T : V \to \mathbb{R}$ and $T(\mathbf{v}_1) = 1$, $T(\mathbf{v}_2) = -1$, find $T(3\mathbf{v}_1 - 5\mathbf{v}_2)$.

b. If $T : V \to \mathbb{R}$ and $T(\mathbf{v}_1) = 2$, $T(\mathbf{v}_2) = -3$, find $T(3\mathbf{v}_1 + 2\mathbf{v}_2)$.

c. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $T\left[\begin{array}{c} 1 \\ 3 \end{array}\right] = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$, $T\left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, find $T\left[\begin{array}{r} -1 \\ 3 \end{array}\right]$.

d. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $T\left[\begin{array}{r} 1 \\ -1 \end{array}\right] = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, $T\left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$, find $T\left[\begin{array}{r} 1 \\ -7 \end{array}\right]$.

e. If $T : \mathbf{P}_2 \to \mathbf{P}_2$ and $T(x+1) = x$, $T(x-1) = 1$, $T(x^2) = 0$, find $T(2 + 3x - x^2)$.

f. If $T : \mathbf{P}_2 \to \mathbb{R}$ and $T(x+2) = 1$, $T(1) = 5$, $T(x^2 + x) = 0$, find $T(2 - x + 3x^2)$.

Exercise 7.1.4 In each case, find a linear transformation with the given properties and compute $T(\mathbf{v})$.

a. $T : \mathbb{R}^2 \to \mathbb{R}^3$; $T(1, 2) = (1, 0, 1)$, $T(-1, 0) = (0, 1, 1)$; $\mathbf{v} = (2, 1)$

b. $T : \mathbb{R}^2 \to \mathbb{R}^3$; $T(2, -1) = (1, -1, 1)$, $T(1, 1) = (0, 1, 0)$; $\mathbf{v} = (-1, 2)$

c. $T : \mathbf{P}_2 \to \mathbf{P}_3$; $T(x^2) = x^3$, $T(x+1) = 0$, $T(x-1) = x$; $\mathbf{v} = x^2 + x + 1$

d. $T : \mathbf{M}_{22} \to \mathbb{R}$; $T\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right] = 3$, $T\left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right] = -1$, $T\left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right] = 0 = T\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$; $\mathbf{v} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$

Exercise 7.1.5 If $T : V \to V$ is a linear transformation, find $T(\mathbf{v})$ and $T(\mathbf{w})$ if:

a. $T(\mathbf{v} + \mathbf{w}) = \mathbf{v} - 2\mathbf{w}$ and $T(2\mathbf{v} - \mathbf{w}) = 2\mathbf{v}$

b. $T(\mathbf{v} + 2\mathbf{w}) = 3\mathbf{v} - \mathbf{w}$ and $T(\mathbf{v} - \mathbf{w}) = 2\mathbf{v} - 4\mathbf{w}$

Exercise 7.1.6 If $T : V \to W$ is a linear transformation, show that $T(\mathbf{v} - \mathbf{v}_1) = T(\mathbf{v}) - T(\mathbf{v}_1)$ for all $\mathbf{v}$ and $\mathbf{v}_1$ in V.

Exercise 7.1.7 Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis of $\mathbb{R}^2$. Is it possible to have a linear transformation T such that $T(\mathbf{e}_1)$ lies in $\mathbb{R}$ while $T(\mathbf{e}_2)$ lies in $\mathbb{R}^2$? Explain your answer.

Exercise 7.1.8 Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V and let $T : V \to V$ be a linear transformation.

a. If $T(\mathbf{v}_i) = \mathbf{v}_i$ for each i, show that $T = 1_V$.

b. If $T(\mathbf{v}_i) = -\mathbf{v}_i$ for each i, show that $T = -1$ is the scalar operator (see Example 7.1.1).

Exercise 7.1.9 If A is an $m \times n$ matrix, let $C_k(A)$ denote column k of A. Show that $C_k : \mathbf{M}_{mn} \to \mathbb{R}^m$ is a linear transformation for each $k = 1, \dots, n$.

Exercise 7.1.10 Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis of $\mathbb{R}^n$. Given k, $1 \leq k \leq n$, define $P_k : \mathbb{R}^n \to \mathbb{R}^n$ by $P_k(r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n) = r_k\mathbf{e}_k$. Show that P_k a linear transformation for each k.

Exercise 7.1.11 Let $S : V \to W$ and $T : V \to W$ be linear transformations. Given a in $\mathbb{R}$, define functions $(S + T) : V \to W$ and $(aT) : V \to W$ by $(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$ and $(aT)(\mathbf{v}) = aT(\mathbf{v})$ for all $\mathbf{v}$ in V. Show that $S + T$ and aT are linear transformations.

Exercise 7.1.12 Describe all linear transformations $T : \mathbb{R} \to V$.

Exercise 7.1.13 Let V and W be vector spaces, let V be finite dimensional, and let $\mathbf{v} \neq \mathbf{0}$ in V. Given any $\mathbf{w}$ in W, show that there exists a linear transformation $T : V \to W$ with $T(\mathbf{v}) = \mathbf{w}$. [*Hint*: Theorem 6.4.1 and Theorem 7.1.3.]

Exercise 7.1.14 Given $\mathbf{y}$ in $\mathbb{R}^n$, define $S_\mathbf{y} : \mathbb{R}^n \to \mathbb{R}$ by $S_\mathbf{y}(\mathbf{x}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ (where $\cdot$ is the dot product introduced in Section 5.3).

a. Show that $S_\mathbf{y} : \mathbb{R}^n \to \mathbb{R}$ is a linear transformation for any $\mathbf{y}$ in $\mathbb{R}^n$.

b. Show that every linear transformation $T : \mathbb{R}^n \to \mathbb{R}$ arises in this way; that is, $T = S_\mathbf{y}$ for some $\mathbf{y}$ in $\mathbb{R}^n$. [*Hint*: If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, write $S_\mathbf{y}(\mathbf{e}_i) = y_i$ for each i. Use Theorem 7.1.1.]

Exercise 7.1.15 Let $T : V \to W$ be a linear transformation.

a. If U is a subspace of V, show that $T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \text{ in } U\}$ is a subspace of W (called the **image** of U under T).

b. If P is a subspace of W, show that $\{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) \text{ in } P\}$ is a subspace of V (called the **preimage** of P under T).

Exercise 7.1.16 Show that differentiation is the only linear transformation $\mathbf{P}_n \to \mathbf{P}_n$ that satisfies $T(x^k) = kx^{k-1}$ for each $k = 0, 1, 2, \dots, n$.

Exercise 7.1.17 Let $T : V \to W$ be a linear transformation and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ denote vectors in V.

a. If $\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\}$ is linearly independent, show that $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is also independent.

b. Find $T : \mathbb{R}^2 \to \mathbb{R}^2$ for which the converse of part (a) is false.

Exercise 7.1.18 Suppose $T : V \to V$ is a linear operator with the property that $T[T(\mathbf{v})] = \mathbf{v}$ for all $\mathbf{v}$ in V. (For example, transposition in $\mathbf{M}_{nn}$ or conjugation in $\mathbb{C}$.) If $\mathbf{v} \neq \mathbf{0}$ in V, show that $\{\mathbf{v}, T(\mathbf{v})\}$ is linearly independent if and only if $T(\mathbf{v}) \neq \mathbf{v}$ and $T(\mathbf{v}) \neq -\mathbf{v}$.

Exercise 7.1.19 If a and b are real numbers, define $T_{a, b} : \mathbb{C} \to \mathbb{C}$ by $T_{a, b}(r + si) = ra + sbi$ for all $r + si$ in $\mathbb{C}$.

a. Show that $T_{a, b}$ is linear and $T_{a, b}(\overline{z}) = \overline{T_{a, b}(z)}$ for all z in $\mathbb{C}$. (Here $\overline{z}$ denotes the conjugate of z.)

b. If $T : \mathbb{C} \to \mathbb{C}$ is linear and $T(\overline{z}) = \overline{T(z)}$ for all z in $\mathbb{C}$, show that $T = T_{a, b}$ for some real a and b.

Exercise 7.1.20 Show that the following conditions are equivalent for a linear transformation $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$.

1. $\text{tr}\,[T(A)] = \text{tr}\,A$ for all A in $\mathbf{M}_{22}$.

2. $T\left[\begin{array}{cc} r_{11} & r_{12} \\ r_{21} & r_{22} \end{array}\right] = r_{11}B_{11} + r_{12}B_{12} + r_{21}B_{21} + r_{22}B_{22}$ for matrices B_{ij} such that $\text{tr}\,B_{11} = 1 = \text{tr}\,B_{22}$ and $\text{tr}\,B_{12} = 0 = \text{tr}\,B_{21}$.

Exercise 7.1.21 Given a in $\mathbb{R}$, consider the **evaluation** map $E_a : \mathbf{P}_n \to \mathbb{R}$ defined in Example 7.1.3.

a. Show that E_a is a linear transformation satisfying the additional condition that $E_a(x^k) = [E_a(x)]^k$ holds for all $k = 0, 1, 2, \dots$. [*Note*: $x^0 = 1$.]

b. If $T : \mathbf{P}_n \to \mathbb{R}$ is a linear transformation satisfying $T(x^k) = [T(x)]^k$ for all $k = 0, 1, 2, \dots$, show that $T = E_a$ for some a in R.

Exercise 7.1.22 If $T : \mathbf{M}_{nn} \to \mathbb{R}$ is any linear transformation satisfying $T(AB) = T(BA)$ for all A and B in $\mathbf{M}_{nn}$, show that there exists a number k such that $T(A) = k \operatorname{tr} A$ for all A. (See Lemma 5.5.1.) [*Hint*: Let E_{ij} denote the $n \times n$ matrix with 1 in the (i, j) position and zeros elsewhere.

Show that $E_{ik}E_{lj} = \begin{cases} 0 & \text{if } k \neq l \\ E_{ij} & \text{if } k = l \end{cases}$. Use this to show that $T(E_{ij}) = 0$ if $i \neq j$ and $T(E_{11}) = T(E_{22}) = \cdots = T(E_{nn})$. Put $k = T(E_{11})$ and use the fact that $\{E_{ij} \mid 1 \leq i, j \leq n\}$ is a basis of $\mathbf{M}_{nn}$.]

Exercise 7.1.23 Let $T : \mathbb{C} \to \mathbb{C}$ be a linear transformation of the real vector space $\mathbb{C}$ and assume that $T(a) = a$ for every real number a. Show that the following are equivalent:

a. $T(zw) = T(z)T(w)$ for all z and w in $\mathbb{C}$.

b. Either $T = 1_{\mathbb{C}}$ or $T(z) = \overline{z}$ for each z in $\mathbb{C}$ (where $\overline{z}$ denotes the conjugate).

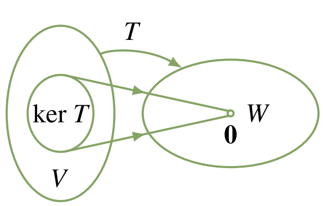
7.2 Kernel and Image of a Linear Transformation

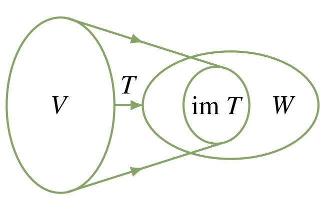
This section is devoted to two important subspaces associated with a linear transformation $T : V \to W$.

Definition 7.2 Kernel and Image of a Linear Transformation

*The **kernel** of T (denoted $\ker T$) and the **image** of T (denoted $\operatorname{im} T$ or $T(V)$) are defined by*

$$\begin{aligned} \ker T &= \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \mathbf{0}\} \\ \operatorname{im} T &= \{T(\mathbf{v}) \mid \mathbf{v} \text{ in } V\} = T(V) \end{aligned}$$





The kernel of T is often called the **nullspace** of T because it consists of all vectors $\mathbf{v}$ in V satisfying the *condition* that $T(\mathbf{v}) = \mathbf{0}$. The image of T is often called the **range** of T and consists of all vectors $\mathbf{w}$ in W of the *form* $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V. These subspaces are depicted in the diagrams.

Example 7.2.1

Let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation induced by the $m \times n$ matrix A, that is $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. Then

$$\begin{aligned} \ker T_A &= \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\} = \operatorname{null} A \quad \text{and} \\ \operatorname{im} T_A &= \{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n\} = \operatorname{im} A \end{aligned}$$

Hence the following theorem extends Example 5.1.2.

Theorem 7.2.1

Let $T : V \to W$ be a linear transformation.

1. $\ker T$ *is a subspace of* V.
2. $\operatorname{im} T$ *is a subspace of* W.

Proof. The fact that $T(\mathbf{0}) = \mathbf{0}$ shows that $\ker T$ and $\operatorname{im} T$ contain the zero vector of V and W respectively.

1. If $\mathbf{v}$ and $\mathbf{v}_1$ lie in $\ker T$, then $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{v}_1)$, so

$$\begin{aligned} T(\mathbf{v} + \mathbf{v}_1) &= T(\mathbf{v}) + T(\mathbf{v}_1) = \mathbf{0} + \mathbf{0} = \mathbf{0} \\ T(r\mathbf{v}) &= rT(\mathbf{v}) = r\mathbf{0} = \mathbf{0} \quad \text{for all } r \text{ in } \mathbb{R} \end{aligned}$$

   Hence $\mathbf{v} + \mathbf{v}_1$ and $r\mathbf{v}$ lie in $\ker T$ (they satisfy the required condition), so $\ker T$ is a subspace of V by the subspace test (Theorem 6.2.1).

2. If $\mathbf{w}$ and $\mathbf{w}_1$ lie in $\operatorname{im} T$, write $\mathbf{w} = T(\mathbf{v})$ and $\mathbf{w}_1 = T(\mathbf{v}_1)$ where $\mathbf{v},\ \mathbf{v}_1 \in V$. Then

$$\begin{aligned} \mathbf{w} + \mathbf{w}_1 &= T(\mathbf{v}) + T(\mathbf{v}_1) = T(\mathbf{v} + \mathbf{v}_1) \\ r\mathbf{w} &= rT(\mathbf{v}) = T(r\mathbf{v}) \quad \text{for all } r \text{ in } \mathbb{R} \end{aligned}$$

   Hence $\mathbf{w} + \mathbf{w}_1$ and $r\mathbf{w}$ both lie in $\operatorname{im} T$ (they have the required form), so $\operatorname{im} T$ is a subspace of W.

□

Given a linear transformation $T : V \to W$:

$\dim(\ker T)$ is called the **nullity** of T and denoted as $\operatorname{nullity}(T)$
$\dim(\operatorname{im} T)$ is called the **rank** of T and denoted as $\operatorname{rank}(T)$

The rank of a matrix A was defined earlier to be the dimension of $\operatorname{col} A$, the column space of A. The two usages of the word *rank* are consistent in the following sense. Recall the definition of T_A in Example 7.2.1.

Example 7.2.2

Given an $m \times n$ matrix A, show that $\operatorname{im} T_A = \operatorname{col} A$, so $\operatorname{rank} T_A = \operatorname{rank} A$.

Solution. Write $A = \left[\begin{array}{ccc} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{array} \right]$ in terms of its columns. Then

$$\operatorname{im} T_A = \{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n\} = \{x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n \mid x_i \text{ in } \mathbb{R}\}$$

using Definition 2.5. Hence $\operatorname{im} T_A$ is the column space of A; the rest follows.

Often, a useful way to study a subspace of a vector space is to exhibit it as the kernel or image of a linear transformation. Here is an example.

Example 7.2.3

Define a transformation $P : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ by $P(A) = A - A^T$ for all A in $\mathbf{M}_{nn}$. Show that P is linear and that:

a. $\ker P$ consists of all symmetric matrices.

b. $\operatorname{im} P$ consists of all skew-symmetric matrices.

Solution. The verification that P is linear is left to the reader. To prove part (a), note that a matrix A lies in $\ker P$ just when $0 = P(A) = A - A^T$, and this occurs if and only if $A = A^T$—that is, A is symmetric. Turning to part (b), the space $\operatorname{im} P$ consists of all matrices $P(A)$, A in $\mathbf{M}_{nn}$. Every such matrix is skew-symmetric because

$$P(A)^T = (A - A^T)^T = A^T - A = -P(A)$$

On the other hand, if S is skew-symmetric (that is, $S^T = -S$), then S lies in $\operatorname{im} P$. In fact,

$$P\left[\tfrac{1}{2}S\right] = \tfrac{1}{2}S - \left[\tfrac{1}{2}S\right]^T = \tfrac{1}{2}(S - S^T) = \tfrac{1}{2}(S + S) = S$$

One-to-One and Onto Transformations

Definition 7.3 One-to-one and Onto Linear Transformations

Let $T : V \to W$ be a linear transformation.

1. *T is said to be* **onto** *if* $\operatorname{im} T = W$.
2. *T is said to be* **one-to-one** *if $T(\mathbf{v}) = T(\mathbf{v}_1)$ implies $\mathbf{v} = \mathbf{v}_1$.*

A vector $\mathbf{w}$ in W is said to be **hit** by T if $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V. Then T is onto if every vector in W is hit at least once, and T is one-to-one if no element of W gets hit twice. Clearly the onto transformations T are those for which $\operatorname{im} T = W$ is as large a subspace of W as possible. By contrast, Theorem 7.2.2 shows that the one-to-one transformations T are the ones with $\ker T$ as *small* a subspace of V as possible.

Theorem 7.2.2

If $T : V \to W$ is a linear transformation, then T is one-to-one if and only if $\ker T = \{\mathbf{0}\}$.

Proof. If T is one-to-one, let $\mathbf{v}$ be any vector in $\ker T$. Then $T(\mathbf{v}) = \mathbf{0}$, so $T(\mathbf{v}) = T(\mathbf{0})$. Hence $\mathbf{v} = \mathbf{0}$ because T is one-to-one. Hence $\ker T = \{\mathbf{0}\}$.

Conversely, assume that $\ker T = \{\mathbf{0}\}$ and let $T(\mathbf{v}) = T(\mathbf{v}_1)$ with $\mathbf{v}$ and $\mathbf{v}_1$ in V. Then $T(\mathbf{v} - \mathbf{v}_1) = T(\mathbf{v}) - T(\mathbf{v}_1) = \mathbf{0}$, so $\mathbf{v} - \mathbf{v}_1$ lies in $\ker T = \{\mathbf{0}\}$. This means that $\mathbf{v} - \mathbf{v}_1 = \mathbf{0}$, so $\mathbf{v} = \mathbf{v}_1$, proving that T is one-to-one. □

Example 7.2.4

The identity transformation $1_V : V \to V$ is both one-to-one and onto for any vector space V.

Example 7.2.5

Consider the linear transformations

$$\begin{aligned} S : \mathbb{R}^3 \to \mathbb{R}^2 \quad &\text{given by } S(x,\ y,\ z) = (x+y,\ x-y) \\ T : \mathbb{R}^2 \to \mathbb{R}^3 \quad &\text{given by } T(x,\ y) = (x+y,\ x-y,\ x) \end{aligned}$$

Show that T is one-to-one but not onto, whereas S is onto but not one-to-one.

Solution. The verification that they are linear is omitted. T is one-to-one because

$$\ker T = \{(x,\ y) \mid x+y = x-y = x = 0\} = \{(0,\ 0)\}$$

However, it is not onto. For example $(0,\ 0,\ 1)$ does not lie in $\operatorname{im} T$ because if $(0,\ 0,\ 1) = (x+y,\ x-y,\ x)$ for some x and y, then $x+y = 0 = x-y$ and $x = 1$, an impossibility. Turning to S, it is not one-to-one by Theorem 7.2.2 because $(0,\ 0,\ 1)$ lies in $\ker S$. But every element $(s,\ t)$ in $\mathbb{R}^2$ lies in $\operatorname{im} S$ because $(s,\ t) = (x+y,\ x-y) = S(x,\ y,\ z)$ for some x, y, and z (in fact, $x = \frac{1}{2}(s+t)$, $y = \frac{1}{2}(s-t)$, and $z = 0$). Hence S is onto.

Example 7.2.6

Let U be an invertible $m \times m$ matrix and define

$$T : \mathbf{M}_{mn} \to \mathbf{M}_{mn} \quad \text{by} \quad T(X) = UX \text{ for all } X \text{ in } \mathbf{M}_{mn}$$

Show that T is a linear transformation that is both one-to-one and onto.

Solution. The verification that T is linear is left to the reader. To see that T is one-to-one, let $T(X) = 0$. Then $UX = 0$, so left-multiplication by U^{-1} gives $X = 0$. Hence $\ker T = \{\mathbf{0}\}$, so T is one-to-one. Finally, if Y is any member of $\mathbf{M}_{mn}$, then $U^{-1}Y$ lies in $\mathbf{M}_{mn}$ too, and $T(U^{-1}Y) = U(U^{-1}Y) = Y$. This shows that T is onto.

The linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ all have the form T_A for some $m \times n$ matrix A (Theorem 2.6.2). The next theorem gives conditions under which they are onto or one-to-one. Note the connection with Theorem 5.4.3 and Theorem 5.4.4.

Theorem 7.2.3

Let A be an $m \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation induced by A, that is $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$.

1. *T_A is onto if and only if* $\operatorname{rank} A = m$.
2. *T_A is one-to-one if and only if* $\operatorname{rank} A = n$.

Proof.

1. We have that $\operatorname{im} T_A$ is the column space of A (see Example 7.2.2), so T_A is onto if and only if the column space of A is $\mathbb{R}^m$. Because the rank of A is the dimension of the column space, this holds if and only if $\operatorname{rank} A = m$.
2. $\ker T_A = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$, so (using Theorem 7.2.2) T_A is one-to-one if and only if $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. This is equivalent to $\operatorname{rank} A = n$ by Theorem 5.4.3. □

The Dimension Theorem

Let A denote an $m \times n$ matrix of rank r and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ denote the corresponding matrix transformation given by $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. It follows from Example 7.2.1 and Example 7.2.2 that $\operatorname{im} T_A = \operatorname{col} A$, so $\dim(\operatorname{im} T_A) = \dim(\operatorname{col} A) = r$. On the other hand Theorem 5.4.2 shows that $\dim(\ker T_A) = \dim(\operatorname{null} A) = n - r$. Combining these we see that

$$\dim(\operatorname{im} T_A) + \dim(\ker T_A) = n \quad \text{for every } m \times n \text{ matrix } A$$

The main result of this section is a deep generalization of this observation.

Theorem 7.2.4: Dimension Theorem

Let $T : V \to W$ be any linear transformation and assume that $\ker T$ and $\operatorname{im} T$ are both finite dimensional. Then V is also finite dimensional and

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$$

In other words, $\dim V = \operatorname{nullity}(T) + \operatorname{rank}(T)$.

Proof. Every vector in $\operatorname{im} T = T(V)$ has the form $T(\mathbf{v})$ for some $\mathbf{v}$ in V. Hence let $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_r)\}$ be a basis of $\operatorname{im} T$, where the $\mathbf{e}_i$ lie in V. Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\}$ be any basis of $\ker T$. Then $\dim(\operatorname{im} T) = r$ and $\dim(\ker T) = k$, so it suffices to show that $B = \{\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{f}_1, \dots, \mathbf{f}_k\}$ is a basis of V.

1. *B spans V.* If $\mathbf{v}$ lies in V, then $T(\mathbf{v})$ lies in $\operatorname{im} V$, so

$$T(\mathbf{v}) = t_1T(\mathbf{e}_1) + t_2T(\mathbf{e}_2) + \cdots + t_rT(\mathbf{e}_r) \quad t_i \text{ in } \mathbb{R}$$

   This implies that $\mathbf{v} - t_1\mathbf{e}_1 - t_2\mathbf{e}_2 - \cdots - t_r\mathbf{e}_r$ lies in $\ker T$ and so is a linear combination of $\mathbf{f}_1, \dots, \mathbf{f}_k$. Hence $\mathbf{v}$ is a linear combination of the vectors in B.

2. B is *linearly independent*. Suppose that t_i and s_j in $\mathbb{R}$ satisfy

$$t_1\mathbf{e}_1 + \cdots + t_r\mathbf{e}_r + s_1\mathbf{f}_1 + \cdots + s_k\mathbf{f}_k = \mathbf{0} \tag{7.1}$$

Applying T gives $t_1T(\mathbf{e}_1) + \cdots + t_rT(\mathbf{e}_r) = \mathbf{0}$ (because $T(\mathbf{f}_i) = \mathbf{0}$ for each i). Hence the independence of $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ yields $t_1 = \cdots = t_r = 0$. But then (7.1) becomes

$$s_1\mathbf{f}_1 + \cdots + s_k\mathbf{f}_k = \mathbf{0}$$

so $s_1 = \cdots = s_k = 0$ by the independence of $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$. This proves that B is linearly independent.

□

Note that the vector space V is not assumed to be finite dimensional in Theorem 7.2.4. In fact, verifying that $\ker T$ and $\operatorname{im} T$ are both finite dimensional is often an important way to *prove* that V is finite dimensional.

Note further that $r + k = n$ in the proof so, after relabelling, we end up with a basis

$$B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$$

of V with the property that $\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ is a basis of $\ker T$ and $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ is a basis of $\operatorname{im} T$. In fact, if V is known in advance to be finite dimensional, then *any* basis $\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ of $\ker T$ can be extended to a basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ of V by Theorem 6.4.1. Moreover, it turns out that, no matter how this is done, the vectors $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ will be a basis of $\operatorname{im} T$. This result is useful, and we record it for reference. The proof is much like that of Theorem 7.2.4 and is left as Exercise 7.2.26.

Theorem 7.2.5

Let $T : V \to W$ be a linear transformation, and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ be a basis of V such that $\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ is a basis of $\ker T$. Then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ is a basis of $\operatorname{im} T$, and hence $r = \operatorname{rank} T$.

The dimension theorem is one of the most useful results in all of linear algebra. It shows that if either $\dim(\ker T)$ or $\dim(\operatorname{im} T)$ can be found, then the other is automatically known. In many cases it is easier to compute one than the other, so the theorem is a real asset. The rest of this section is devoted to illustrations of this fact. The next example uses the dimension theorem to give a different proof of the first part of Theorem 5.4.2.

Example 7.2.7

Let A be an $m \times n$ matrix of rank r. Show that the space $\operatorname{null} A$ of all solutions of the system $A\mathbf{x} = \mathbf{0}$ of m homogeneous equations in n variables has dimension $n - r$.

Solution. The space in question is just $\ker T_A$, where $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is defined by $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. But $\dim(\operatorname{im} T_A) = \operatorname{rank} T_A = \operatorname{rank} A = r$ by Example 7.2.2, so $\dim(\ker T_A) = n - r$ by the dimension theorem.

Example 7.2.8

If $T : V \to W$ is a linear transformation where V is finite dimensional, then

$$\dim(\ker T) \leq \dim V \quad \text{and} \quad \dim(\operatorname{im} T) \leq \dim V$$

Indeed, $\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$ by Theorem 7.2.4. Of course, the first inequality also follows because $\ker T$ is a subspace of V.

Example 7.2.9

Let $D : \mathbf{P}_n \to \mathbf{P}_{n-1}$ be the differentiation map defined by $D[p(x)] = p'(x)$. Compute $\ker D$ and hence conclude that D is onto.

Solution. Because $p'(x) = 0$ means $p(x)$ is constant, we have $\dim(\ker D) = 1$. Since $\dim \mathbf{P}_n = n+1$, the dimension theorem gives

$$\dim(\operatorname{im} D) = (n+1) - \dim(\ker D) = n = \dim(\mathbf{P}_{n-1})$$

This implies that $\operatorname{im} D = \mathbf{P}_{n-1}$, so D is onto.

Of course it is not difficult to verify directly that each polynomial $q(x)$ in $\mathbf{P}_{n-1}$ is the derivative of some polynomial in $\mathbf{P}_n$ (simply integrate $q(x)$!), so the dimension theorem is not needed in this case. However, in some situations it is difficult to see directly that a linear transformation is onto, and the method used in Example 7.2.9 may be by far the easiest way to prove it. Here is another illustration.

Example 7.2.10

Given a in $\mathbb{R}$, the evaluation map $E_a : \mathbf{P}_n \to \mathbb{R}$ is given by $E_a[p(x)] = p(a)$. Show that E_a is linear and onto, and hence conclude that $\{(x-a), (x-a)^2, \dots, (x-a)^n\}$ is a basis of $\ker E_a$, the subspace of all polynomials $p(x)$ for which $p(a) = 0$.

Solution. E_a is linear by Example 7.1.3; the verification that it is onto is left to the reader. Hence $\dim(\operatorname{im} E_a) = \dim(\mathbb{R}) = 1$, so $\dim(\ker E_a) = (n+1) - 1 = n$ by the dimension theorem. Now each of the n polynomials $(x-a), (x-a)^2, \dots, (x-a)^n$ clearly lies in $\ker E_a$, and they are linearly independent (they have distinct degrees). Hence they are a basis because $\dim(\ker E_a) = n$.

We conclude by applying the dimension theorem to the rank of a matrix.

Example 7.2.11

If A is any $m \times n$ matrix, show that $\operatorname{rank} A = \operatorname{rank} A^T A = \operatorname{rank} AA^T$.

Solution. It suffices to show that $\operatorname{rank} A = \operatorname{rank} A^T A$ (the rest follows by replacing A with A^T). Write $B = A^T A$, and consider the associated matrix transformations

$$T_A : \mathbb{R}^n \to \mathbb{R}^m \quad \text{and} \quad T_B : \mathbb{R}^n \to \mathbb{R}^n$$

The dimension theorem and Example 7.2.2 give

$$\begin{aligned}\text{rank } A &= \text{rank } T_A = \dim(\text{im } T_A) = n - \dim(\ker T_A)\\ \text{rank } B &= \text{rank } T_B = \dim(\text{im } T_B) = n - \dim(\ker T_B)\end{aligned}$$

so it suffices to show that $\ker T_A = \ker T_B$. Now $A\mathbf{x} = \mathbf{0}$ implies that $B\mathbf{x} = A^T A\mathbf{x} = \mathbf{0}$, so $\ker T_A$ is contained in $\ker T_B$. On the other hand, if $B\mathbf{x} = \mathbf{0}$, then $A^T A\mathbf{x} = \mathbf{0}$, so

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$$

This implies that $A\mathbf{x} = \mathbf{0}$, so $\ker T_B$ is contained in $\ker T_A$.

Exercises for 7.2

Exercise 7.2.1 For each matrix A, find a basis for the kernel and image of T_A, and find the rank and nullity of T_A.

a. $\left[\begin{array}{rrrr} 1 & 2 & -1 & 1 \\ 3 & 1 & 0 & 2 \\ 1 & -3 & 2 & 0 \end{array}\right]$ b. $\left[\begin{array}{rrrr} 2 & 1 & -1 & 3 \\ 1 & 0 & 3 & 1 \\ 1 & 1 & -4 & 2 \end{array}\right]$

c. $\left[\begin{array}{rrr} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 4 & -1 & 5 \\ 0 & 2 & -2 \end{array}\right]$ d. $\left[\begin{array}{rrr} 2 & 1 & 0 \\ 1 & -1 & 3 \\ 1 & 2 & -3 \\ 0 & 3 & -6 \end{array}\right]$

Exercise 7.2.2 In each case, (i) find a basis of $\ker T$, and (ii) find a basis of $\text{im } T$. You may assume that T is linear.

a. $T : \mathbf{P}_2 \to \mathbb{R}^2$; $T(a + bx + cx^2) = (a, b)$

b. $T : \mathbf{P}_2 \to \mathbb{R}^2$; $T(p(x)) = (p(0), p(1))$

c. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x + y, x + y, 0)$

d. $T : \mathbb{R}^3 \to \mathbb{R}^4$; $T(x, y, z) = (x, x, y, y)$

e. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} a+b & b+c \\ c+d & d+a \end{array}\right]$

f. $T : \mathbf{M}_{22} \to \mathbb{R}$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = a + d$

g. $T : \mathbf{P}_n \to \mathbb{R}$; $T(r_0 + r_1x + \cdots + r_nx^n) = r_n$

h. $T : \mathbb{R}^n \to \mathbb{R}$; $T(r_1, r_2, \dots, r_n) = r_1 + r_2 + \cdots + r_n$

i. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T(X) = XA - AX$, where $A = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$

j. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T(X) = XA$, where $A = \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]$

Exercise 7.2.3 Let $P : V \to \mathbb{R}$ and $Q : V \to \mathbb{R}$ be linear transformations, where V is a vector space. Define $T : V \to \mathbb{R}^2$ by $T(\mathbf{v}) = (P(\mathbf{v}), Q(\mathbf{v}))$.

a. Show that T is a linear transformation.

b. Show that $\ker T = \ker P \cap \ker Q$, the set of vectors in both $\ker P$ and $\ker Q$.

Exercise 7.2.4 In each case, find a basis $B = \{\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ of V such that $\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ is a basis of $\ker T$, and verify Theorem 7.2.5.

a. $T : \mathbb{R}^3 \to \mathbb{R}^4$; $T(x, y, z) = (x - y + 2z, x + y - z, 2x + z, 2y - 3z)$

b. $T : \mathbb{R}^3 \to \mathbb{R}^4$; $T(x, y, z) = (x + y + z, 2x - y + 3z, z - 3y, 3x + 4z)$

Exercise 7.2.5 Show that every matrix X in $\mathbf{M}_{nn}$ has the form $X = A^T - 2A$ for some matrix A in $\mathbf{M}_{nn}$. [*Hint*: The dimension theorem.]

Exercise 7.2.6 In each case either prove the statement or give an example in which it is false. Throughout, let $T : V \to W$ be a linear transformation where V and W are finite dimensional.

a. If $V = W$, then $\ker T \subseteq \operatorname{im} T$.

b. If $\dim V = 5$, $\dim W = 3$, and $\dim(\ker T) = 2$, then T is onto.

c. If $\dim V = 5$ and $\dim W = 4$, then $\ker T \neq \{\mathbf{0}\}$.

d. If $\ker T = V$, then $W = \{\mathbf{0}\}$.

e. If $W = \{\mathbf{0}\}$, then $\ker T = V$.

f. If $W = V$, and $\operatorname{im} T \subseteq \ker T$, then $T = 0$.

g. If $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a basis of V and $T(\mathbf{e}_1) = \mathbf{0} = T(\mathbf{e}_2)$, then $\dim(\operatorname{im} T) \leq 1$.

h. If $\dim(\ker T) \leq \dim W$, then $\dim W \geq \frac{1}{2} \dim V$.

i. If T is one-to-one, then $\dim V \leq \dim W$.

j. If $\dim V \leq \dim W$, then T is one-to-one.

k. If T is onto, then $\dim V \geq \dim W$.

l. If $\dim V \geq \dim W$, then T is onto.

m. If $\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_k)\}$ is independent, then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is independent.

n. If $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ spans V, then $\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_k)\}$ spans W.

Exercise 7.2.7 Show that linear independence is preserved by one-to-one transformations and that spanning sets are preserved by onto transformations. More precisely, if $T : V \to W$ is a linear transformation, show that:

a. If T is one-to-one and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is independent in V, then $\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\}$ is independent in W.

b. If T is onto and $V = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, then $W = \operatorname{span}\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\}$.

Exercise 7.2.8 Given $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ in a vector space V, define $T : \mathbb{R}^n \to V$ by $T(r_1, \dots, r_n) = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$. Show that T is linear, and that:

a. T is one-to-one if and only if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is independent.

b. T is onto if and only if $V = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Exercise 7.2.9 Let $T : V \to V$ be a linear transformation where V is finite dimensional. Show that exactly one of (i) and (ii) holds: (i) $T(\mathbf{v}) = \mathbf{0}$ for some $\mathbf{v} \neq \mathbf{0}$ in V; (ii) $T(\mathbf{x}) = \mathbf{v}$ has a solution $\mathbf{x}$ in V for every $\mathbf{v}$ in V.

Exercise 7.2.10 Let $T : \mathbf{M}_{nn} \to \mathbb{R}$ denote the trace map: $T(A) = \operatorname{tr} A$ for all A in $\mathbf{M}_{nn}$. Show that $\dim(\ker T) = n^2 - 1$.

Exercise 7.2.11 Show that the following are equivalent for a linear transformation $T : V \to W$.

1. $\ker T = V$
2. $\operatorname{im} T = \{\mathbf{0}\}$
3. $T = 0$

Exercise 7.2.12 Let A and B be $m \times n$ and $k \times n$ matrices, respectively. Assume that $A\mathbf{x} = \mathbf{0}$ implies $B\mathbf{x} = \mathbf{0}$ for every n-column $\mathbf{x}$. Show that $\operatorname{rank} A \geq \operatorname{rank} B$.
[*Hint*: Theorem 7.2.4.]

Exercise 7.2.13 Let A be an $m \times n$ matrix of rank r. Thinking of $\mathbb{R}^n$ as rows, define $V = \{\mathbf{x} \text{ in } \mathbb{R}^m \mid \mathbf{x}A = \mathbf{0}\}$. Show that $\dim V = m - r$.

Exercise 7.2.14 Consider

$$V = \left\{ \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \middle| a + c = b + d \right\}$$

a. Consider $S : \mathbf{M}_{22} \to \mathbb{R}$ with $S \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = a + c - b - d$. Show that S is linear and onto and that V is a subspace of $\mathbf{M}_{22}$. Compute $\dim V$.

b. Consider $T : V \to \mathbb{R}$ with $T \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = a + c$. Show that T is linear and onto, and use this information to compute $\dim(\ker T)$.

Exercise 7.2.15 Define $T : \mathbf{P}_n \to \mathbb{R}$ by $T[p(x)] =$ the sum of all the coefficients of $p(x)$.

a. Use the dimension theorem to show that $\dim(\ker T) = n$.

b. Conclude that $\{x-1,\ x^2-1,\ \dots,\ x^n-1\}$ is a basis of $\ker T$.

Exercise 7.2.16 Use the dimension theorem to prove Theorem 1.3.1: If A is an $m \times n$ matrix with $m < n$, the system $A\mathbf{x} = \mathbf{0}$ of m homogeneous equations in n variables always has a nontrivial solution.

Exercise 7.2.17 Let B be an $n \times n$ matrix, and consider the subspaces $U = \{A \mid A \text{ in } \mathbf{M}_{mn},\ AB = 0\}$ and $V = \{AB \mid A \text{ in } \mathbf{M}_{mn}\}$. Show that $\dim U + \dim V = mn$.

Exercise 7.2.18 Let U and V denote, respectively, the spaces of even and odd polynomials in $\mathbf{P}_n$. Show that $\dim U + \dim V = n+1$. [*Hint*: Consider $T : \mathbf{P}_n \to \mathbf{P}_n$ where $T[p(x)] = p(x) - p(-x)$.]

Exercise 7.2.19 Show that every polynomial $f(x)$ in $\mathbf{P}_{n-1}$ can be written as $f(x) = p(x+1) - p(x)$ for some polynomial $p(x)$ in $\mathbf{P}_n$. [*Hint*: Define $T : \mathbf{P}_n \to \mathbf{P}_{n-1}$ by $T[p(x)] = p(x+1) - p(x)$.]

Exercise 7.2.20 Let U and V denote the spaces of symmetric and skew-symmetric $n \times n$ matrices. Show that $\dim U + \dim V = n^2$.

Exercise 7.2.21 Assume that B in $\mathbf{M}_{nn}$ satisfies $B^k = 0$ for some $k \geq 1$. Show that every matrix in $\mathbf{M}_{nn}$ has the form $BA - A$ for some A in $\mathbf{M}_{nn}$. [*Hint*: Show that $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ is linear and one-to-one where $T(A) = BA - A$ for each A.]

Exercise 7.2.22 Fix a column $\mathbf{y} \neq \mathbf{0}$ in $\mathbb{R}^n$ and let $U = \{A \text{ in } \mathbf{M}_{nn} \mid A\mathbf{y} = \mathbf{0}\}$. Show that $\dim U = n(n-1)$.

Exercise 7.2.23 If B in $\mathbf{M}_{mn}$ has rank r, let $U = \{A \text{ in } \mathbf{M}_{nn} \mid BA = 0\}$ and $W = \{BA \mid A \text{ in } \mathbf{M}_{nn}\}$. Show that $\dim U = n(n-r)$ and $\dim W = nr$. [*Hint*: Show that U consists of all matrices A whose columns are in the null space of B. Use Example 7.2.7.]

Exercise 7.2.24 Let $T : V \to V$ be a linear transformation where $\dim V = n$. If $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$, show that every vector $\mathbf{v}$ in V can be written $\mathbf{v} = \mathbf{u} + \mathbf{w}$ for some $\mathbf{u}$ in $\ker T$ and $\mathbf{w}$ in $\operatorname{im} T$. [*Hint*: Choose bases $B \subseteq \ker T$ and $D \subseteq \operatorname{im} T$, and use Exercise 6.3.33.]

Exercise 7.2.25 Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator of rank 1, where $\mathbb{R}^n$ is written as rows. Show that there exist numbers $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ such that $T(X) = XA$ for all rows X in $\mathbb{R}^n$, where

$$A = \left[\begin{array}{cccc} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{array}\right]$$

[*Hint*: $\operatorname{im} T = \mathbb{R}\mathbf{w}$ for $\mathbf{w} = (b_1, \dots, b_n)$ in $\mathbb{R}^n$.]

Exercise 7.2.26 Prove Theorem 7.2.5.

Exercise 7.2.27 Let $T : V \to \mathbb{R}$ be a nonzero linear transformation, where $\dim V = n$. Show that there is a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of V so that $T(r_1\mathbf{e}_1 + r_2\mathbf{e}_2 + \cdots + r_n\mathbf{e}_n) = r_1$.

Exercise 7.2.28 Let $f \neq 0$ be a fixed polynomial of degree $m \geq 1$. If p is any polynomial, recall that $(p \circ f)(x) = p[f(x)]$. Define $T_f : P_n \to P_{n+m}$ by $T_f(p) = p \circ f$.

a. Show that T_f is linear.

b. Show that T_f is one-to-one.

Exercise 7.2.29 Let U be a subspace of a finite dimensional vector space V.

a. Show that $U = \ker T$ for some linear operator $T : V \to V$.

b. Show that $U = \operatorname{im} S$ for some linear operator $S : V \to V$. [*Hint*: Theorem 6.4.1 and Theorem 7.1.3.]

Exercise 7.2.30 Let V and W be finite dimensional vector spaces.

a. Show that $\dim W \leq \dim V$ if and only if there exists an onto linear transformation $T : V \to W$. [*Hint*: Theorem 6.4.1 and Theorem 7.1.3.]

b. Show that $\dim W \geq \dim V$ if and only if there exists a one-to-one linear transformation $T : V \to W$. [*Hint*: Theorem 6.4.1 and Theorem 7.1.3.]

Exercise 7.2.31 Let A and B be $n \times n$ matrices, and assume that $AXB = 0$, $X \in \mathbf{M}_{nn}$, implies $X = 0$. Show that A and B are both invertible. [Hint: Dimension Theorem.]

7.3 Isomorphisms and Composition

Often two vector spaces can consist of quite different types of vectors but, on closer examination, turn out to be the same underlying space displayed in different symbols. For example, consider the spaces

$$\mathbb{R}^2 = \{(a,\ b) \mid a,\ b \in \mathbb{R}\} \quad \text{and} \quad \mathbf{P}_1 = \{a + bx \mid a,\ b \in \mathbb{R}\}$$

Compare the addition and scalar multiplication in these spaces:

$$\begin{aligned} (a,\ b) + (a_1,\ b_1) &= (a + a_1,\ b + b_1) \\ r(a,\ b) &= (ra,\ rb) \end{aligned} \qquad \begin{aligned} (a + bx) + (a_1 + b_1x) &= (a + a_1) + (b + b_1)x \\ r(a + bx) &= (ra) + (rb)x \end{aligned}$$

Clearly these are the *same* vector space expressed in different notation: if we change each $(a,\ b)$ in $\mathbb{R}^2$ to $a + bx$, then $\mathbb{R}^2$ *becomes* $\mathbf{P}_1$, complete with addition and scalar multiplication. This can be expressed by noting that the map $(a,\ b) \mapsto a + bx$ is a linear transformation $\mathbb{R}^2 \to \mathbf{P}_1$ that is both one-to-one and onto. In this form, we can describe the general situation.

Definition 7.4 Isomorphic Vector Spaces

A linear transformation $T : V \to W$ is called an ***isomorphism*** *if it is both onto and one-to-one. The vector spaces V and W are said to be* ***isomorphic*** *if there exists an isomorphism $T : V \to W$, and we write $V \cong W$ when this is the case.*

Example 7.3.1

The identity transformation $1_V : V \to V$ is an isomorphism for any vector space V.

Example 7.3.2

If $T : \mathbf{M}_{mn} \to \mathbf{M}_{nm}$ is defined by $T(A) = A^T$ for all A in $\mathbf{M}_{mn}$, then T is an isomorphism (verify). Hence $\mathbf{M}_{mn} \cong \mathbf{M}_{nm}$.

Example 7.3.3

Isomorphic spaces can "look" quite different. For example, $\mathbf{M}_{22} \cong \mathbf{P}_3$ because the map $T : \mathbf{M}_{22} \to \mathbf{P}_3$ given by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + bx + cx^2 + dx^3$ is an isomorphism (verify).

The word *isomorphism* comes from two Greek roots: *iso*, meaning "same," and *morphos*, meaning "form." An isomorphism $T : V \to W$ induces a pairing

$$\mathbf{v} \leftrightarrow T(\mathbf{v})$$

between vectors $\mathbf{v}$ in V and vectors $T(\mathbf{v})$ in W that preserves vector addition and scalar multiplication. Hence, *as far as their vector space properties are concerned*, the spaces V and W are identical except for notation. Because addition and scalar multiplication in either space are completely determined by the same operations in the other space, all *vector space* properties of either space are completely determined by those of the other.

One of the most important examples of isomorphic spaces was considered in Chapter 4. Let A denote the set of all "arrows" with tail at the origin in space, and make A into a vector space using the parallelogram law and the scalar multiple law (see Section 4.1). Then define a transformation $T : \mathbb{R}^3 \to A$ by taking

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \text{the arrow } \mathbf{v} \text{ from the origin to the point } P(x,\ y,\ z).$$

In Section 4.1 matrix addition and scalar multiplication were shown to correspond to the parallelogram law and the scalar multiplication law for these arrows, so the map T is a linear transformation. Moreover T is an isomorphism: it is one-to-one by Theorem 4.1.2, and it is onto because, given an arrow $\mathbf{v}$ in A with tip $P(x,\ y,\ z)$, we have $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{v}$. This justifies the identification $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in Chapter 4 of the geometric arrows with the algebraic matrices. This identification is very useful. The arrows give a "picture" of the matrices and so bring geometric intuition into $\mathbb{R}^3$; the matrices are useful for detailed calculations and so bring analytic precision into geometry. This is one of the best examples of the power of an isomorphism to shed light on both spaces being considered.

The following theorem gives a very useful characterization of isomorphisms: They are the linear transformations that preserve bases.

Theorem 7.3.1

If V and W are finite dimensional spaces, the following conditions are equivalent for a linear transformation $T : V \to W$.

1. *T is an isomorphism.*
2. *If $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_n\}$ is any basis of V, then $\{T(\mathbf{e}_1),\ T(\mathbf{e}_2),\ \dots,\ T(\mathbf{e}_n)\}$ is a basis of W.*
3. *There exists a basis $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_n\}$ of V such that $\{T(\mathbf{e}_1),\ T(\mathbf{e}_2),\ \dots,\ T(\mathbf{e}_n)\}$ is a basis of W.*

Proof. (1) $\Rightarrow$ (2). Let $\{\mathbf{e}_1,\ \dots,\ \mathbf{e}_n\}$ be a basis of V. If $t_1T(\mathbf{e}_1)+\cdots+t_nT(\mathbf{e}_n) = \mathbf{0}$ with t_i in $\mathbb{R}$, then $T(t_1\mathbf{e}_1+\cdots+t_n\mathbf{e}_n) = \mathbf{0}$, so $t_1\mathbf{e}_1+\cdots+t_n\mathbf{e}_n = \mathbf{0}$ (because $\ker T = \{\mathbf{0}\}$). But then each $t_i = 0$ by the independence of the $\mathbf{e}_i$, so $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_n)\}$ is independent. To show that it spans W, choose $\mathbf{w}$ in W. Because T is onto, $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V, so write $\mathbf{v} = t_1\mathbf{e}_1+\cdots+t_n\mathbf{e}_n$. Hence we obtain $\mathbf{w} = T(\mathbf{v}) = t_1T(\mathbf{e}_1)+\cdots+t_nT(\mathbf{e}_n)$, proving that $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_n)\}$ spans W.

(2) $\Rightarrow$ (3). This is because V has a basis.

(3) $\Rightarrow$ (1). If $T(\mathbf{v}) = \mathbf{0}$, write $\mathbf{v} = v_1\mathbf{e}_1+\cdots+v_n\mathbf{e}_n$ where each v_i is in $\mathbb{R}$. Then

$$\mathbf{0} = T(\mathbf{v}) = v_1T(\mathbf{e}_1)+\cdots+v_nT(\mathbf{e}_n)$$

so $v_1 = \cdots = v_n = 0$ by (3). Hence $\mathbf{v} = \mathbf{0}$, so $\ker T = \{\mathbf{0}\}$ and T is one-to-one. To show that T is onto, let $\mathbf{w}$ be any vector in W. By (3) there exist $w_1, \dots, w_n$ in $\mathbb{R}$ such that

$$\mathbf{w} = w_1 T(\mathbf{e}_1) + \cdots + w_n T(\mathbf{e}_n) = T(w_1\mathbf{e}_1 + \cdots + w_n\mathbf{e}_n)$$

Thus T is onto. □

Theorem 7.3.1 dovetails nicely with Theorem 7.1.3 as follows. Let V and W be vector spaces of dimension n, and suppose that $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ are bases of V and W, respectively. Theorem 7.1.3 asserts that there exists a linear transformation $T : V \to W$ such that

$$T(\mathbf{e}_i) = \mathbf{f}_i \quad \text{for each } i = 1, 2, \dots, n$$

Then $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$ is evidently a basis of W, so T is an isomorphism by Theorem 7.3.1. Furthermore, the action of T is prescribed by

$$T(r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n) = r_1\mathbf{f}_1 + \cdots + r_n\mathbf{f}_n$$

so isomorphisms between spaces of equal dimension can be easily defined as soon as bases are known. In particular, this shows that if two vector spaces V and W have the same dimension then they are isomorphic, that is $V \cong W$. This is half of the following theorem.

Theorem 7.3.2

If V and W are finite dimensional vector spaces, then $V \cong W$ if and only if $\dim V = \dim W$.

Proof. It remains to show that if $V \cong W$ then $\dim V = \dim W$. But if $V \cong W$, then there exists an isomorphism $T : V \to W$. Since V is finite dimensional, let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis of V. Then $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$ is a basis of W by Theorem 7.3.1, so $\dim W = n = \dim V$. □

Corollary 7.3.1

Let U, V, and W denote vector spaces. Then:

1. *$V \cong V$ for every vector space V.*
2. *If $V \cong W$ then $W \cong V$.*
3. *If $U \cong V$ and $V \cong W$, then $U \cong W$.*

The proof is left to the reader. By virtue of these properties, the relation $\cong$ is called an *equivalence relation* on the class of finite dimensional vector spaces. Since $\dim(\mathbb{R}^n) = n$ it follows that

Corollary 7.3.2

If V is a vector space and $\dim V = n$, then V is isomorphic to $\mathbb{R}^n$.

If V is a vector space of dimension n, note that there are important explicit isomorphisms $V \to \mathbb{R}^n$. Fix a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of V and write $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ for the standard basis of $\mathbb{R}^n$. By Theorem 7.1.3 there is a unique linear transformation $C_B : V \to \mathbb{R}^n$ given by

$$C_B(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n = \left[\begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array}\right]$$

where each v_i is in $\mathbb{R}$. Moreover, $C_B(\mathbf{b}_i) = \mathbf{e}_i$ for each i so C_B is an isomorphism by Theorem 7.3.1, called the **coordinate isomorphism** corresponding to the basis B. These isomorphisms will play a central role in Chapter 9.

The conclusion in the above corollary can be phrased as follows: As far as vector space properties are concerned, every n-dimensional vector space V is essentially the same as $\mathbb{R}^n$; they are the "same" vector space except for a change of symbols. This appears to make the process of abstraction seem less important—just study $\mathbb{R}^n$ and be done with it! But consider the different "feel" of the spaces $\mathbf{P}_8$ and $\mathbf{M}_{33}$ even though they are both the "same" as $\mathbb{R}^9$: For example, vectors in $\mathbf{P}_8$ can have roots, while vectors in $\mathbf{M}_{33}$ can be multiplied. So the merit in the abstraction process lies in identifying *common* properties of the vector spaces in the various examples. This is important even for finite dimensional spaces. However, the payoff from abstraction is much greater in the infinite dimensional case, particularly for spaces of functions.

Example 7.3.4

Let V denote the space of all 2×2 symmetric matrices. Find an isomorphism $T : \mathbf{P}_2 \to V$ such that $T(1) = I$, where I is the 2×2 identity matrix.

Solution. $\{1, x, x^2\}$ is a basis of $\mathbf{P}_2$, and we want a basis of V containing I. The set $\left\{ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right] \right\}$ is independent in V, so it is a basis because $\dim V = 3$ (by Example 6.3.11). Hence define $T : \mathbf{P}_2 \to V$ by taking $T(1) = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$, $T(x) = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$, $T(x^2) = \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$, and extending linearly as in Theorem 7.1.3. Then T is an isomorphism by Theorem 7.3.1, and its action is given by

$$T(a + bx + cx^2) = aT(1) + bT(x) + cT(x^2) = \left[\begin{array}{cc} a & b \\ b & a+c \end{array}\right]$$

The dimension theorem (Theorem 7.2.4) gives the following useful fact about isomorphisms.

Theorem 7.3.3

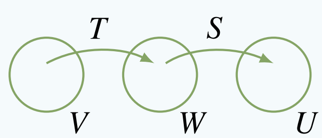
If V and W have the same dimension n, a linear transformation $T : V \to W$ is an isomorphism if it is either one-to-one or onto.

Proof. The dimension theorem asserts that $\dim(\ker T) + \dim(\operatorname{im} T) = n$, so $\dim(\ker T) = 0$ if and only if $\dim(\operatorname{im} T) = n$. Thus T is one-to-one if and only if T is onto, and the result follows. □

Composition

Suppose that $T : V \to W$ and $S : W \to U$ are linear transformations. They link together as in the diagram so, as in Section 2.3, it is possible to define a new function $V \to U$ by first applying T and then S.

Definition 7.5 Composition of Linear Transformations



Given linear transformations $V \xrightarrow{T} W \xrightarrow{S} U$, the ***composite*** *$ST : V \to U$ of T and S is defined by*

$$ST(\mathbf{v}) = S[T(\mathbf{v})] \quad \text{for all } \mathbf{v} \text{ in } V$$

The operation of forming the new function ST is called ***composition***.[1]

The action of ST can be described compactly as follows: ST means first T then S.

Not all pairs of linear transformations can be composed. For example, if $T : V \to W$ and $S : W \to U$ are linear transformations then $ST : V \to U$ is defined, but TS *cannot* be formed unless $U = V$ because $S : W \to U$ and $T : V \to W$ do not "link" in that order.[2]

Moreover, even if ST and TS can both be formed, they may not be equal. In fact, if $S : \mathbb{R}^m \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ are induced by matrices A and B respectively, then ST and TS can both be formed (they are induced by AB and BA respectively), but the matrix products AB and BA may not be equal (they may not even be the same *size*). Here is another example.

Example 7.3.5

Define: $S : \mathbf{M}_{22} \to \mathbf{M}_{22}$ and $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ by $S\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$ and $T(A) = A^T$ for $A \in \mathbf{M}_{22}$. Describe the action of ST and TS, and show that $ST \neq TS$.

Solution. $ST\begin{bmatrix} a & b \\ c & d \end{bmatrix} = S\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} b & d \\ a & c \end{bmatrix}$, whereas
$TS\begin{bmatrix} a & b \\ c & d \end{bmatrix} = T\begin{bmatrix} c & d \\ a & b \end{bmatrix} = \begin{bmatrix} c & a \\ d & b \end{bmatrix}$.
It is clear that $TS\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ need not equal $ST\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, so $TS \neq ST$.

The next theorem collects some basic properties of the composition operation.

[1]In Section 2.3 we denoted the composite as $S \circ T$. However, it is more convenient to use the simpler notation ST.
[2]Actually, all that is required is $U \subseteq V$.

Theorem 7.3.4: [3]

Let $V \xrightarrow{T} W \xrightarrow{S} U \xrightarrow{R} Z$ be linear transformations.

1. *The composite ST is again a linear transformation.*
2. *$T1_V = T$ and $1_W T = T$.*
3. *$(RS)T = R(ST)$.*

Proof. The proofs of (1) and (2) are left as Exercise 7.3.25. To prove (3), observe that, for all $\mathbf{v}$ in V:

$$\{(RS)T\}(\mathbf{v}) = (RS)\left[T(\mathbf{v})\right] = R\{S\left[T(\mathbf{v})\right]\} = R\{(ST)(\mathbf{v})\} = \{R(ST)\}(\mathbf{v})$$

□

Up to this point, composition seems to have no connection with isomorphisms. In fact, the two notions are closely related.

Theorem 7.3.5

Let V and W be finite dimensional vector spaces. The following conditions are equivalent for a linear transformation $T : V \to W$.

1. *T is an isomorphism.*
2. *There exists a linear transformation $S : W \to V$ such that $ST = 1_V$ and $TS = 1_W$.*

Moreover, in this case S is also an isomorphism and is uniquely determined by T:

$$\text{If } \mathbf{w} \text{ in } W \text{ is written as } \mathbf{w} = T(\mathbf{v}), \quad \text{then } S(\mathbf{w}) = \mathbf{v}.$$

Proof. (1) $\Rightarrow$ (2). If $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis of V, then $D = \{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$ is a basis of W by Theorem 7.3.1. Hence (using Theorem 7.1.3), define a linear transformation $S : W \to V$ by

$$S[T(\mathbf{e}_i)] = \mathbf{e}_i \quad \text{for each } i \tag{7.2}$$

Since $\mathbf{e}_i = 1_V(\mathbf{e}_i)$, this gives $ST = 1_V$ by Theorem 7.1.2. But applying T gives $T\left[S\left[T(\mathbf{e}_i)\right]\right] = T(\mathbf{e}_i)$ for each i, so $TS = 1_W$ (again by Theorem 7.1.2, using the basis D of W).

(2) $\Rightarrow$ (1). If $T(\mathbf{v}) = T(\mathbf{v}_1)$, then $S\left[T(\mathbf{v})\right] = S\left[T(\mathbf{v}_1)\right]$. Because $ST = 1_V$ by (2), this reads $\mathbf{v} = \mathbf{v}_1$; that is, T is one-to-one. Given $\mathbf{w}$ in W, the fact that $TS = 1_W$ means that $\mathbf{w} = T\left[S(\mathbf{w})\right]$, so T is onto.

[3]Theorem 7.3.4 can be expressed by saying that vector spaces and linear transformations are an example of a category. In general a category consists of certain objects and, for any two objects X and Y, a set $\operatorname{mor}(X, Y)$. The elements α of $\operatorname{mor}(X, Y)$ are called morphisms from X to Y and are written $\alpha : X \to Y$. It is assumed that identity morphisms and composition are defined in such a way that Theorem 7.3.4 holds. Hence, in the category of vector spaces the objects are the vector spaces themselves and the morphisms are the linear transformations. Another example is the category of metric spaces, in which the objects are sets equipped with a distance function (called a metric), and the morphisms are continuous functions (with respect to the metric). The category of sets and functions is a very basic example.

Finally, S is uniquely determined by the condition $ST = 1_V$ because this condition implies (7.2). S is an isomorphism because it carries the basis D to B. As to the last assertion, given $\mathbf{w}$ in W, write $\mathbf{w} = r_1T(\mathbf{e}_1) + \cdots + r_nT(\mathbf{e}_n)$. Then $\mathbf{w} = T(\mathbf{v})$, where $\mathbf{v} = r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n$. Then $S(\mathbf{w}) = \mathbf{v}$ by (7.2). $\square$

Given an isomorphism $T : V \to W$, the unique isomorphism $S : W \to V$ satisfying condition (2) of Theorem 7.3.5 is called the **inverse** of T and is denoted by T^{-1}. Hence $T : V \to W$ and $T^{-1} : W \to V$ are related by the **fundamental identities**:

$$T^{-1}[T(\mathbf{v})] = \mathbf{v} \text{ for all } \mathbf{v} \text{ in } V \quad \text{and} \quad T\left[T^{-1}(\mathbf{w})\right] = \mathbf{w} \text{ for all } \mathbf{w} \text{ in } W$$

In other words, each of T and T^{-1} reverses the action of the other. In particular, equation (7.2) in the proof of Theorem 7.3.5 shows how to define T^{-1} using the image of a basis under the isomorphism T. Here is an example.

Example 7.3.6

Define $T : \mathbf{P}_1 \to \mathbf{P}_1$ by $T(a+bx) = (a-b)+ax$. Show that T has an inverse, and find the action of T^{-1}.

Solution. The transformation T is linear (verify). Because $T(1) = 1+x$ and $T(x) = -1$, T carries the basis $B = \{1, x\}$ to the basis $D = \{1+x, -1\}$. Hence T is an isomorphism, and T^{-1} carries D back to B, that is,

$$T^{-1}(1+x) = 1 \quad \text{and} \quad T^{-1}(-1) = x$$

Because $a+bx = b(1+x)+(b-a)(-1)$, we obtain

$$T^{-1}(a+bx) = bT^{-1}(1+x)+(b-a)T^{-1}(-1) = b+(b-a)x$$

Sometimes the action of the inverse of a transformation is apparent.

Example 7.3.7

If $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ is a basis of a vector space V, the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism defined by

$$C_B(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = (v_1, v_2, \dots, v_n)^T$$

The way to reverse the action of C_B is clear: $C_B^{-1} : \mathbb{R}^n \to V$ is given by

$$C_B^{-1}(v_1, v_2, \dots, v_n) = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \quad \text{for all } v_i \text{ in } V$$

Condition (2) in Theorem 7.3.5 characterizes the inverse of a linear transformation $T : V \to W$ as the (unique) transformation $S : W \to V$ that satisfies $ST = 1_V$ and $TS = 1_W$. This often determines the inverse.

Example 7.3.8

Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(x, y, z) = (z, x, y)$. Show that $T^3 = 1_{\mathbb{R}^3}$, and hence find T^{-1}.

Solution. $T^2(x, y, z) = T[T(x, y, z)] = T(z, x, y) = (y, z, x)$. Hence

$$T^3(x, y, z) = T\left[T^2(x, y, z)\right] = T(y, z, x) = (x, y, z)$$

Since this holds for all (x, y, z), it shows that $T^3 = 1_{\mathbb{R}^3}$, so $T(T^2) = 1_{\mathbb{R}^3} = (T^2)T$. Thus $T^{-1} = T^2$ by (2) of Theorem 7.3.5.

Example 7.3.9

Define $T : \mathbf{P}_n \to \mathbb{R}^{n+1}$ by $T(p) = (p(0), p(1), \dots, p(n))$ for all p in $\mathbf{P}_n$. Show that T^{-1} exists.

Solution. The verification that T is linear is left to the reader. If $T(p) = 0$, then $p(k) = 0$ for $k = 0, 1, \dots, n$, so p has $n+1$ distinct roots. Because p has degree at most n, this implies that $p = 0$ is the zero polynomial (Theorem 6.5.4) and hence that T is one-to-one. But $\dim \mathbf{P}_n = n+1 = \dim \mathbb{R}^{n+1}$, so this means that T is also onto and hence is an isomorphism. Thus T^{-1} exists by Theorem 7.3.5. Note that we have not given a description of the action of T^{-1}, we have merely shown that such a description exists. To give it explicitly requires some ingenuity; one method involves the Lagrange interpolation expansion (Theorem 6.5.3).

Exercises for 7.3

Exercise 7.3.1 Verify that each of the following is an isomorphism (Theorem 7.3.3 is useful).

a. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x+y, y+z, z+x)$

b. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x, x+y, x+y+z)$

c. $T : \mathbb{C} \to \mathbb{C}$; $T(z) = \overline{z}$

d. $T : \mathbf{M}_{mn} \to \mathbf{M}_{mn}$; $T(X) = UXV$, U and V invertible

e. $T : \mathbf{P}_1 \to \mathbb{R}^2$; $T[p(x)] = [p(0), p(1)]$

f. $T : V \to V$; $T(\mathbf{v}) = k\mathbf{v}$, $k \neq 0$ a fixed number, V any vector space

g. $T : \mathbf{M}_{22} \to \mathbb{R}^4$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = (a+b, d, c, a-b)$

h. $T : \mathbf{M}_{mn} \to \mathbf{M}_{nm}$; $T(A) = A^T$

Exercise 7.3.2 Show that

$$\{a+bx+cx^2, a_1+b_1x+c_1x^2, a_2+b_2x+c_2x^2\}$$

is a basis of $\mathbf{P}_2$ if and only if $\{(a, b, c), (a_1, b_1, c_1), (a_2, b_2, c_2)\}$ is a basis of $\mathbb{R}^3$.

Exercise 7.3.3 If V is any vector space, let V^n denote the space of all n-tuples $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$, where each $\mathbf{v}_i$ lies in V. (This is a vector space with component-wise operations; see Exercise 6.1.17.) If $C_j(A)$ denotes the jth column of the $m \times n$ matrix A, show that $T : \mathbf{M}_{mn} \to (\mathbb{R}^m)^n$ is an isomorphism if $T(A) = \left[\begin{array}{cccc} C_1(A) & C_2(A) & \cdots & C_n(A) \end{array}\right]$. (Here $\mathbb{R}^m$ consists of columns.)

Exercise 7.3.4 In each case, compute the action of ST and TS, and show that $ST \neq TS$.

a. $S : \mathbb{R}^2 \to \mathbb{R}^2$ with $S(x, y) = (y, x)$; $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(x, y) = (x, 0)$

b. $S : \mathbb{R}^3 \to \mathbb{R}^3$ with $S(x, y, z) = (x, 0, z)$; $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $T(x, y, z) = (x+y, 0, y+z)$

c. $S : \mathbf{P}_2 \to \mathbf{P}_2$ with $S(p) = p(0) + p(1)x + p(2)x^2$; $T : \mathbf{P}_2 \to \mathbf{P}_2$ with $T(a+bx+cx^2) = b+cx+ax^2$

d. $S : \mathbf{M}_{22} \to \mathbf{M}_{22}$ with $S\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} a & 0 \\ 0 & d \end{array}\right]$; $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ with $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} c & a \\ d & b \end{array}\right]$

Exercise 7.3.5 In each case, show that the linear transformation T satisfies $T^2 = T$.

a. $T : \mathbb{R}^4 \to \mathbb{R}^4$; $T(x, y, z, w) = (x, 0, z, 0)$

b. $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (x+y, 0)$

c. $T : \mathbf{P}_2 \to \mathbf{P}_2$; $T(a+bx+cx^2) = (a+b-c)+cx+cx^2$

d. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \frac{1}{2}\left[\begin{array}{cc} a+c & b+d \\ a+c & b+d \end{array}\right]$

Exercise 7.3.6 Determine whether each of the following transformations T has an inverse and, if so, determine the action of T^{-1}.

a. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x+y, y+z, z+x)$

b. $T : \mathbb{R}^4 \to \mathbb{R}^4$; $T(x, y, z, t) = (x+y, y+z, z+t, t+x)$

c. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} a-c & b-d \\ 2a-c & 2b-d \end{array}\right]$

d. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} a+2c & b+2d \\ 3c-a & 3d-b \end{array}\right]$

e. $T : \mathbf{P}_2 \to \mathbb{R}^3$; $T(a+bx+cx^2) = (a-c, 2b, a+c)$

f. $T : \mathbf{P}_2 \to \mathbb{R}^3$; $T(p) = [p(0), p(1), p(-1)]$

Exercise 7.3.7 In each case, show that T is self-inverse, that is: $T^{-1} = T$.

a. $T : \mathbb{R}^4 \to \mathbb{R}^4$; $T(x, y, z, w) = (x, -y, -z, w)$

b. $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (ky-x, y)$, k any fixed number

c. $T : \mathbf{P}_n \to \mathbf{P}_n$; $T(p(x)) = p(3-x)$

d. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T(X) = AX$ where $A = \frac{1}{4}\left[\begin{array}{cc} 5 & -3 \\ 3 & -5 \end{array}\right]$

Exercise 7.3.8 In each case, show that $T^6 = 1_{R^4}$ and so determine T^{-1}.

a. $T : \mathbb{R}^4 \to \mathbb{R}^4$; $T(x, y, z, w) = (-x, z, w, y)$

b. $T : \mathbb{R}^4 \to \mathbb{R}^4$; $T(x, y, z, w) = (-y, x-y, z, -w)$

Exercise 7.3.9 In each case, show that T is an isomorphism by defining T^{-1} explicitly.

a. $T : \mathbf{P}_n \to \mathbf{P}_n$ is given by $T[p(x)] = p(x+1)$.

b. $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ is given by $T(A) = UA$ where U is invertible in $\mathbf{M}_{nn}$.

Exercise 7.3.10 Given linear transformations $V \xrightarrow{T} W \xrightarrow{S} U$:

a. If S and T are both one-to-one, show that ST is one-to-one.

b. If S and T are both onto, show that ST is onto.

Exercise 7.3.11 Let $T : V \to W$ be a linear transformation.

a. If T is one-to-one and $TR = TR_1$ for transformations R and $R_1 : U \to V$, show that $R = R_1$.

b. If T is onto and $ST = S_1T$ for transformations S and $S_1 : W \to U$, show that $S = S_1$.

Exercise 7.3.12 Consider the linear transformations $V \xrightarrow{T} W \xrightarrow{R} U$.

a. Show that $\ker T \subseteq \ker RT$.

b. Show that $\operatorname{im} RT \subseteq \operatorname{im} R$.

Exercise 7.3.13 Let $V \xrightarrow{T} U \xrightarrow{S} W$ be linear transformations.

a. If ST is one-to-one, show that T is one-to-one and that $\dim V \leq \dim U$.

b. If ST is onto, show that S is onto and that $\dim W \leq \dim U$.

Exercise 7.3.14 Let $T : V \to V$ be a linear transformation. Show that $T^2 = 1_V$ if and only if T is invertible and $T = T^{-1}$.

Exercise 7.3.15 Let N be a nilpotent $n \times n$ matrix (that is, $N^k = 0$ for some k). Show that $T : \mathbf{M}_{nm} \to \mathbf{M}_{nm}$ is an isomorphism if $T(X) = X - NX$. [*Hint*: If X is in $\ker T$, show that $X = NX = N^2X = \cdots$. Then use Theorem 7.3.3.]

Exercise 7.3.16 Let $T : V \to W$ be a linear transformation, and let $\{\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ be any basis of V such that $\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ is a basis of $\ker T$. Show that $\operatorname{im} T \cong \operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$. [*Hint*: See Theorem 7.2.5.]

Exercise 7.3.17 Is every isomorphism $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ given by an invertible matrix U such that $T(X) = UX$ for all X in $\mathbf{M}_{22}$? Prove your answer.

Exercise 7.3.18 Let $\mathbf{D}_n$ denote the space of all functions f from $\{1, 2, \dots, n\}$ to $\mathbb{R}$ (see Exercise 6.3.35). If $T : \mathbf{D}_n \to \mathbb{R}^n$ is defined by

$$T(f) = (f(1), f(2), \dots, f(n)),$$

show that T is an isomorphism.

Exercise 7.3.19

a. Let V be the vector space of Exercise 6.1.3. Find an isomorphism $T : V \to \mathbb{R}^1$.

b. Let V be the vector space of Exercise 6.1.4. Find an isomorphism $T : V \to \mathbb{R}^2$.

Exercise 7.3.20 Let $V \xrightarrow{T} W \xrightarrow{S} V$ be linear transformations such that $ST = 1_V$. If $\dim V = \dim W = n$, show that $S = T^{-1}$ and $T = S^{-1}$. [*Hint*: Exercise 7.3.13 and Theorem 7.3.3, Theorem 7.3.4, and Theorem 7.3.5.]

Exercise 7.3.21 Let $V \xrightarrow{T} W \xrightarrow{S} V$ be functions such that $TS = 1_W$ and $ST = 1_V$. If T is linear, show that S is also linear.

Exercise 7.3.22 Let A and B be matrices of size $p \times m$ and $n \times q$. Assume that $mn = pq$. Define $R : \mathbf{M}_{mn} \to \mathbf{M}_{pq}$ by $R(X) = AXB$.

a. Show that $\mathbf{M}_{mn} \cong \mathbf{M}_{pq}$ by comparing dimensions.

b. Show that R is a linear transformation.

c. Show that if R is an isomorphism, then $m = p$ and $n = q$. [*Hint*: Show that $T : \mathbf{M}_{mn} \to \mathbf{M}_{pn}$ given by $T(X) = AX$ and $S : \mathbf{M}_{mn} \to \mathbf{M}_{mq}$ given by $S(X) = XB$ are both one-to-one, and use the dimension theorem.]

Exercise 7.3.23 Let $T : V \to V$ be a linear transformation such that $T^2 = 0$ is the zero transformation.

a. If $V \neq \{\mathbf{0}\}$, show that T cannot be invertible.

b. If $R : V \to V$ is defined by $R(\mathbf{v}) = \mathbf{v} + T(\mathbf{v})$ for all $\mathbf{v}$ in V, show that R is linear and invertible.

Exercise 7.3.24 Let V consist of all sequences $[x_0, x_1, x_2, \dots)$ of numbers, and define vector operations

$$[x_o, x_1, \dots) + [y_0, y_1, \dots) = [x_0 + y_0, x_1 + y_1, \dots)$$
$$r[x_0, x_1, \dots) = [rx_0, rx_1, \dots)$$

a. Show that V is a vector space of infinite dimension.

b. Define $T : V \to V$ and $S : V \to V$ by $T[x_0, x_1, \dots) = [x_1, x_2, \dots)$ and $S[x_0, x_1, \dots) = [0, x_0, x_1, \dots)$. Show that $TS = 1_V$, so TS is one-to-one and onto, but that T is not one-to-one and S is not onto.

Exercise 7.3.25 Prove (1) and (2) of Theorem 7.3.4.

Exercise 7.3.26 Define $T : \mathbf{P}_n \to \mathbf{P}_n$ by $T(p) = p(x) + xp'(x)$ for all p in $\mathbf{P}_n$.

a. Show that T is linear.

b. Show that $\ker T = \{\mathbf{0}\}$ and conclude that T is an isomorphism. [*Hint*: Write $p(x) = a_0 + a_1x + \cdots + a_nx^n$ and compare coefficients if $p(x) = -xp'(x)$.]

c. Conclude that each $q(x)$ in $\mathbf{P}_n$ has the form $q(x) = p(x) + xp'(x)$ for some unique polynomial $p(x)$.

d. Does this remain valid if T is defined by $T[p(x)] = p(x) - xp'(x)$? Explain.

Exercise 7.3.27 Let $T : V \to W$ be a linear transformation, where V and W are finite dimensional.

a. Show that T is one-to-one if and only if there exists a linear transformation $S : W \to V$ with $ST = 1_V$. [*Hint*: If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis of V and T is one-to-one, show that W has a basis $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n), \mathbf{f}_{n+1}, \dots, \mathbf{f}_{n+k}\}$ and use Theorem 7.1.2 and Theorem 7.1.3.]

b. Show that T is onto if and only if there exists a linear transformation $S : W \to V$ with $TS = 1_W$. [*Hint*: Let $\{\mathbf{e}_1, \dots, \mathbf{e}_r, \dots, \mathbf{e}_n\}$ be a basis of V such that $\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ is a basis of $\ker T$. Use Theorem 7.2.5, Theorem 7.1.2 and Theorem 7.1.3.]

Exercise 7.3.28 Let S and T be linear transformations $V \to W$, where $\dim V = n$ and $\dim W = m$.

a. Show that $\ker S = \ker T$ if and only if $T = RS$ for some isomorphism $R : W \to W$. [*Hint*: Let $\{\mathbf{e}_1, \dots, \mathbf{e}_r, \dots, \mathbf{e}_n\}$ be a basis of V such that $\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ is a basis of $\ker S = \ker T$. Use Theorem 7.2.5 to extend $\{S(\mathbf{e}_1), \dots, S(\mathbf{e}_r)\}$ and $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_r)\}$ to bases of W.]

b. Show that $\operatorname{im} S = \operatorname{im} T$ if and only if $T = SR$ for some isomorphism $R : V \to V$. [*Hint*: Show that $\dim(\ker S) = \dim(\ker T)$ and choose bases $\{\mathbf{e}_1, \dots, \mathbf{e}_r, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \dots, \mathbf{f}_r, \dots, \mathbf{f}_n\}$ of V where $\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_{r+1}, \dots, \mathbf{f}_n\}$ are bases of $\ker S$ and $\ker T$, respectively. If $1 \le i \le r$, show that $S(\mathbf{e}_i) = T(\mathbf{g}_i)$ for some $\mathbf{g}_i$ in V, and prove that $\{\mathbf{g}_1, \dots, \mathbf{g}_r, \mathbf{f}_{r+1}, \dots, \mathbf{f}_n\}$ is a basis of V.]

Exercise 7.3.29 If $T : V \to V$ is a linear transformation where $\dim V = n$, show that $TST = T$ for some isomorphism $S : V \to V$. [*Hint*: Let $\{\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ be as in Theorem 7.2.5. Extend $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_r)\}$ to a basis of V, and use Theorem 7.3.1, Theorem 7.1.2 and Theorem 7.1.3.]

Exercise 7.3.30 Let A and B denote $m \times n$ matrices. In each case show that (1) and (2) are equivalent.

a. (1) A and B have the same null space. (2) $B = PA$ for some invertible $m \times m$ matrix P.

b. (1) A and B have the same range. (2) $B = AQ$ for some invertible $n \times n$ matrix Q.

[*Hint*: Use Exercise 7.3.28.]

7.4 A Theorem about Differential Equations

Differential equations are instrumental in solving a variety of problems throughout science, social science, and engineering. In this brief section, we will see that the set of solutions of a linear differential equation (with constant coefficients) is a vector space and we will calculate its dimension. The proof is pure linear algebra, although the applications are primarily in analysis. However, a key result (Lemma 7.4.3 below) can be applied much more widely.

We denote the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ by f', and f will be called **differentiable** if it can be differentiated any number of times. If f is a differentiable function, the nth derivative $f^{(n)}$ of f is the result of differentiating n times. Thus $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f^{(1)\prime}$, $\dots$, and in general $f^{(n+1)} = f^{(n)\prime}$ for each $n \ge 0$. For small values of n these are often written as $f, f', f'', f''', \dots$.

If a, b, and c are numbers, the differential equations

$$f'' - af' - bf = 0 \quad \text{or} \quad f''' - af'' - bf' - cf = 0$$

are said to be of **second order** and **third-order**, respectively. In general, an equation

$$f^{(n)} - a_{n-1}f^{(n-1)} - a_{n-2}f^{(n-2)} - \cdots - a_2 f^{(2)} - a_1 f^{(1)} - a_0 f^{(0)} = 0, \; a_i \text{ in } \mathbb{R} \tag{7.3}$$

is called a **differential equation of order** n. We want to describe all solutions of this equation. Of course a knowledge of calculus is required.

The set $\mathbf{F}$ of all functions $\mathbb{R} \to \mathbb{R}$ is a vector space with operations as described in Example 6.1.7. If f and g are differentiable, we have $(f+g)' = f' + g'$ and $(af)' = af'$ for all a in $\mathbb{R}$. With this it is a routine matter to verify that the following set is a subspace of $\mathbf{F}$:

$$\mathbf{D}_n = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is differentiable and is a solution to (7.3)}\}$$

Our sole objective in this section is to prove

Theorem 7.4.1

The space $\mathbf{D}_n$ has dimension n.

As will be clear later, the proof of Theorem 7.4.1 requires that we enlarge $\mathbf{D}_n$ somewhat and allow our differentiable functions to take values in the set $\mathbb{C}$ of complex numbers. To do this, we must clarify what it means for a function $f : \mathbb{R} \to \mathbb{C}$ to be differentiable. For each real number x write $f(x)$ in terms of its real and imaginary parts $f_r(x)$ and $f_i(x)$:

$$f(x) = f_r(x) + i f_i(x)$$

This produces new functions $f_r : \mathbb{R} \to \mathbb{R}$ and $f_i : \mathbb{R} \to \mathbb{R}$, called the **real** and **imaginary parts** of f, respectively. We say that f is **differentiable** if both f_r and f_i are differentiable (as real functions), and we define the **derivative** f' of f by

$$f' = f_r' + i f_i' \tag{7.4}$$

We refer to this frequently in what follows.[4]

With this, write $\mathbf{D}_\infty$ for the set of all differentiable complex valued functions $f : \mathbb{R} \to \mathbb{C}$. This is a *complex* vector space using pointwise addition (see Example 6.1.7), and the following scalar multiplication: For any w in $\mathbb{C}$ and f in $\mathbf{D}_\infty$, we define $wf : \mathbb{R} \to \mathbb{C}$ by $(wf)(x) = wf(x)$ for all x in $\mathbb{R}$. We will be working in $\mathbf{D}_\infty$ for the rest of this section. In particular, consider the following complex subspace of $\mathbf{D}_\infty$:

$$\mathbf{D}_n^* = \{f : \mathbb{R} \to \mathbb{C} \mid f \text{ is a solution to (7.3)}\}$$

Clearly, $\mathbf{D}_n \subseteq \mathbf{D}_n^*$, and our interest in $\mathbf{D}_n^*$ comes from

Lemma 7.4.1

If $\dim_{\mathbb{C}}(\mathbf{D}_n^*) = n$, then $\dim_{\mathbb{R}}(\mathbf{D}_n) = n$.

Proof. Observe first that if $\dim_{\mathbb{C}}(\mathbf{D}_n^*) = n$, then $\dim_{\mathbb{R}}(\mathbf{D}_n^*) = 2n$. [In fact, if $\{g_1, \dots, g_n\}$ is a $\mathbb{C}$-basis of $\mathbf{D}_n^*$ then $\{g_1, \dots, g_n, ig_1, \dots, ig_n\}$ is a $\mathbb{R}$-basis of $\mathbf{D}_n^*$]. Now observe that the set $\mathbf{D}_n \times \mathbf{D}_n$ of all ordered pairs (f, g) with f and g in $\mathbf{D}_n$ is a real vector space with componentwise operations. Define

$$\theta : \mathbf{D}_n^* \to \mathbf{D}_n \times \mathbf{D}_n \quad \text{given by} \quad \theta(f) = (f_r, f_i) \text{ for } f \text{ in } \mathbf{D}_n^*$$

[4]Write $|w|$ for the absolute value of any complex number w. As for functions $\mathbb{R} \to \mathbb{R}$, we say that $\lim_{t\to 0} f(t) = w$ if, for all $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(t) - w| <\in$ whenever $|t| < \delta$. (Note that t represents a real number here.) In particular, given a real number x, we define the *derivative* f' of a function $f : \mathbb{R} \to \mathbb{C}$ by $f'(x) = \lim_{t\to 0} \left\{\frac{1}{t}[f(x+t) - f(x)]\right\}$ and we say that f is *differentiable* if $f'(x)$ exists for all x in $\mathbb{R}$. Then we can *prove* that f is differentiable if and only if both f_r and f_i are differentiable, and that $f' = f_r' + i f_i'$ in this case.

One verifies that θ is onto and one-to-one, and it is $\mathbb{R}$-linear because $f \to f_r$ and $f \to f_i$ are both $\mathbb{R}$-linear. Hence $\mathbf{D}_n^* \cong \mathbf{D}_n \times \mathbf{D}_n$ as $\mathbb{R}$-spaces. Since $\dim_{\mathbb{R}}(\mathbf{D}_n^*)$ is finite, it follows that $\dim_{\mathbb{R}}(\mathbf{D}_n)$ is finite, and we have

$$2\dim_{\mathbb{R}}(\mathbf{D}_n) = \dim_{\mathbb{R}}(\mathbf{D}_n \times \mathbf{D}_n) = \dim_{\mathbb{R}}(\mathbf{D}_n^*) = 2n$$

Hence $\dim_{\mathbb{R}}(\mathbf{D}_n) = n$, as required. □

It follows that to prove Theorem 7.4.1 it suffices to show that $\dim_{\mathbb{C}}(\mathbf{D}_n^*) = n$.

There is one function that arises frequently in any discussion of differential equations. Given a complex number $w = a + ib$ (where a and b are real), we have $e^w = e^a(\cos b + i \sin b)$. The law of exponents, $e^w e^v = e^{w+v}$ for all w, v in $\mathbb{C}$ is easily verified using the formulas for $\sin(b + b_1)$ and $\cos(b + b_1)$. If x is a variable and $w = a + ib$ is a complex number, define the **exponential function** e^{wx} by

$$e^{wx} = e^{ax}(\cos bx + i \sin bx)$$

Hence e^{wx} is differentiable because its real and imaginary parts are differentiable for all x. Moreover, the following can be proved using (7.4):

$$(e^{wx})' = we^{wx}$$

In addition, (7.4) gives the **product rule** for differentiation:

$$\text{If } f \text{ and } g \text{ are in } \mathbf{D}_\infty, \text{ then } (fg)' = f'g + fg'$$

We omit the verifications.

To prove that $\dim_{\mathbb{C}}(\mathbf{D}_n^*) = n$, two preliminary results are required. Here is the first.

Lemma 7.4.2

Given f in $\mathbf{D}_\infty$ and w in $\mathbb{C}$, there exists g in $\mathbf{D}_\infty$ such that $g' - wg = f$.

Proof. Define $p(x) = f(x)e^{-wx}$. Then p is differentiable, whence p_r and p_i are both differentiable, hence continuous, and so both have antiderivatives, say $p_r = q_r'$ and $p_i = q_i'$. Then the function $q = q_r + iq_i$ is in $\mathbf{D}_\infty$, and $q' = p$ by (7.4). Finally define $g(x) = q(x)e^{wx}$. Then

$$g' = q'e^{wx} + qwe^{wx} = pe^{wx} + w(qe^{wx}) = f + wg$$

by the product rule, as required. □

The second preliminary result is important in its own right.

Lemma 7.4.3: Kernel Lemma

Let V be a vector space, and let S and T be linear operators $V \to V$. If S is onto and both $\ker(S)$ and $\ker(T)$ are finite dimensional, then $\ker(TS)$ is also finite dimensional and $\dim[\ker(TS)] = \dim[\ker(T)] + \dim[\ker(S)]$.

Proof. Let $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ be a basis of $\ker(T)$ and let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis of $\ker(S)$. Since S is onto, let $\mathbf{u}_i = S(\mathbf{w}_i)$ for some $\mathbf{w}_i$ in V. It suffices to show that

$$B = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

is a basis of $\ker(TS)$. Note $B \subseteq \ker(TS)$ because $TS(\mathbf{w}_i) = T(\mathbf{u}_i) = \mathbf{0}$ for each i and $TS(\mathbf{v}_j) = T(\mathbf{0}) = \mathbf{0}$ for each j.

Spanning. If $\mathbf{v}$ is in $\ker(TS)$, then $S(\mathbf{v})$ is in $\ker(T)$, say $S(\mathbf{v}) = \sum r_i\mathbf{u}_i = \sum r_i S(\mathbf{w}_i) = S(\sum r_i\mathbf{w}_i)$. It follows that $\mathbf{v} - \sum r_i\mathbf{w}_i$ is in $\ker(S) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, proving that $\mathbf{v}$ is in $\text{span}(B)$.

Independence. Let $\sum r_i\mathbf{w}_i + \sum t_j\mathbf{v}_j = \mathbf{0}$. Applying S, and noting that $S(\mathbf{v}_j) = \mathbf{0}$ for each j, yields $\mathbf{0} = \sum r_i S(\mathbf{w}_i) = \sum r_i\mathbf{u}_i$. Hence $r_i = 0$ for each i, and so $\sum t_j\mathbf{v}_j = \mathbf{0}$. This implies that each $t_j = 0$, and so proves the independence of B. □

Proof of Theorem 7.4.1. By Lemma 7.4.1, it suffices to prove that $\dim_{\mathbb{C}}(\mathbf{D}_n^*) = n$. This holds for $n = 1$ because the proof of Theorem 3.5.1 goes through to show that $\mathbf{D}_1^* = \mathbb{C}e^{a_0x}$. Hence we proceed by induction on n. With an eye on equation (7.3), consider the polynomial

$$p(t) = t^n - a_{n-1}t^{n-1} - a_{n-2}t^{n-2} - \cdots - a_2t^2 - a_1t - a_0$$

(called the *characteristic polynomial* of equation (7.3)). Now define a map $D : \mathbf{D}_\infty \to \mathbf{D}_\infty$ by $D(f) = f'$ for all f in $\mathbf{D}_\infty$. Then D is a linear operator, whence $p(D) : \mathbf{D}_\infty \to \mathbf{D}_\infty$ is also a linear operator. Moreover, since $D^k(f) = f^{(k)}$ for each $k \geq 0$, equation (7.3) takes the form $p(D)(f) = 0$. In other words,

$$\mathbf{D}_n^* = \ker[p(D)]$$

By the fundamental theorem of algebra,[5] let w be a complex root of $p(t)$, so that $p(t) = q(t)(t - w)$ for some complex polynomial $q(t)$ of degree $n - 1$. It follows that $p(D) = q(D)(D - w1_{\mathbf{D}_\infty})$. Moreover $D - w1_{\mathbf{D}_\infty}$ is onto by Lemma 7.4.2, $\dim_{\mathbb{C}}[\ker(D - w1_{\mathbf{D}_\infty})] = 1$ by the case $n = 1$ above, and $\dim_{\mathbb{C}}(\ker[q(D)]) = n - 1$ by induction. Hence Lemma 7.4.3 shows that $\ker[P(D)]$ is also finite dimensional and

$$\dim_{\mathbb{C}}(\ker[p(D)]) = \dim_{\mathbb{C}}(\ker[q(D)]) + \dim_{\mathbb{C}}(\ker[D - w1_{\mathbf{D}_\infty}]) = (n-1) + 1 = n.$$

Since $\mathbf{D}_n^* = \ker[p(D)]$, this completes the induction, and so proves Theorem 7.4.1. □

7.5 More on Linear Recurrences[6]

In Section 3.4 we used diagonalization to study linear recurrences, and gave several examples. We now apply the theory of vector spaces and linear transformations to study the problem in more generality.

Consider the linear recurrence

$$x_{n+2} = 6x_n - x_{n+1} \quad \text{for } n \geq 0$$

If the initial values x_0 and x_1 are prescribed, this gives a sequence of numbers. For example, if $x_0 = 1$ and $x_1 = 1$ the sequence continues

$$x_2 = 5,\ x_3 = 1,\ x_4 = 29,\ x_5 = -23,\ x_6 = 197,\ \dots$$

[5]This is the reason for allowing our solutions to (7.3) to be *complex* valued.

[6]This section requires only Sections 7.1-7.3.

as the reader can verify. Clearly, the entire sequence is uniquely determined by the recurrence and the two initial values. In this section we define a vector space structure on the set of *all* sequences, and study the subspace of those sequences that satisfy a particular recurrence.

Sequences will be considered entities in their own right, so it is useful to have a special notation for them. Let

$$[x_n) \quad \text{denote the sequence } x_0,\ x_1,\ x_2,\ \dots,\ x_n,\ \dots$$

Example 7.5.1

$$\begin{array}{ll} [n) & \text{is the sequence } 0,\ 1,\ 2,\ 3,\ \dots \\ [n+1) & \text{is the sequence } 1,\ 2,\ 3,\ 4,\ \dots \\ [2^n) & \text{is the sequence } 1,\ 2,\ 2^2,\ 2^3,\ \dots \\ [(-1)^n) & \text{is the sequence } 1,\ -1,\ 1,\ -1,\ \dots \\ [5) & \text{is the sequence } 5,\ 5,\ 5,\ 5,\ \dots \end{array}$$

Sequences of the form $[c)$ for a fixed number c will be referred to as **constant sequences**, and those of the form $[\lambda^n)$, λ some number, are **power sequences**.

Two sequences are regarded as **equal** when they are identical:

$$[x_n) = [y_n) \quad \text{means} \quad x_n = y_n \quad \text{for all } n = 0,\ 1,\ 2,\ \dots$$

Addition and scalar multiplication of sequences are defined by

$$\begin{aligned} [x_n) + [y_n) &= [x_n + y_n) \\ r[x_n) &= [rx_n) \end{aligned}$$

These operations are analogous to the addition and scalar multiplication in $\mathbb{R}^n$, and it is easy to check that the vector-space axioms are satisfied. The zero vector is the constant sequence $[0)$, and the negative of a sequence $[x_n)$ is given by $-[x_n) = [-x_n)$.

Now suppose k real numbers $r_0,\ r_1,\ \dots,\ r_{k-1}$ are given, and consider the **linear recurrence relation** determined by these numbers.

$$x_{n+k} = r_0x_n + r_1x_{n+1} + \cdots + r_{k-1}x_{n+k-1} \tag{7.5}$$

When $r_0 \neq 0$, we say this recurrence has **length** k.[7] For example, the relation $x_{n+2} = 2x_n + x_{n+1}$ is of length 2.

A sequence $[x_n)$ is said to **satisfy** the relation (7.5) if (7.5) holds for all $n \geq 0$. Let V denote the set of all sequences that satisfy the relation. In symbols,

$$V = \{[x_n) \mid x_{n+k} = r_0x_n + r_1x_{n+1} + \cdots + r_{k-1}x_{n+k-1} \text{ hold for all } n \geq 0\}$$

It is easy to see that the constant sequence $[0)$ lies in V and that V is closed under addition and scalar multiplication of sequences. Hence V is vector space (being a subspace of the space of all sequences). The following important observation about V is needed (it was used implicitly earlier): If the first k terms of two sequences agree, then the sequences are identical. More formally,

[7] We shall usually assume that $r_0 \neq 0$; otherwise, we are essentially dealing with a recurrence of shorter length than k.

Lemma 7.5.1

Let $[x_n)$ and $[y_n)$ denote two sequences in V. Then

$$[x_n) = [y_n) \quad \textit{if and only if} \quad x_0 = y_0,\ x_1 = y_1,\ \dots,\ x_{k-1} = y_{k-1}$$

Proof. If $[x_n) = [y_n)$ then $x_n = y_n$ for *all* $n = 0,\ 1,\ 2,\ \dots$. Conversely, if $x_i = y_i$ for all $i = 0,\ 1,\ \dots,\ k-1$, use the recurrence (7.5) for $n = 0$.

$$x_k = r_0x_0 + r_1x_1 + \cdots + r_{k-1}x_{k-1} = r_0y_0 + r_1y_1 + \cdots + r_{k-1}y_{k-1} = y_k$$

Next the recurrence for $n = 1$ establishes $x_{k+1} = y_{k+1}$. The process continues to show that $x_{n+k} = y_{n+k}$ holds for *all* $n \geq 0$ by induction on n. Hence $[x_n) = [y_n)$. □

This shows that a sequence in V is completely determined by its first k terms. In particular, given a k-tuple $\mathbf{v} = (v_0,\ v_1,\ \dots,\ v_{k-1})$ in $\mathbb{R}^k$, define

$$T(\mathbf{v}) \text{ to be the sequence in } V \text{ whose first } k \text{ terms are } v_0,\ v_1,\ \dots,\ v_{k-1}$$

The rest of the sequence $T(\mathbf{v})$ is determined by the recurrence, so $T : \mathbb{R}^k \to V$ is a function. In fact, it is an isomorphism.

Theorem 7.5.1

Given real numbers $r_0,\ r_1,\ \dots,\ r_{k-1}$, let

$$V = \{[x_n) \mid x_{n+k} = r_0x_n + r_1x_{n+1} + \cdots + r_{k-1}x_{n+k-1}, \ \textit{for all } n \geq 0\}$$

denote the vector space of all sequences satisfying the linear recurrence relation (7.5) determined by $r_0,\ r_1,\ \dots,\ r_{k-1}$. Then the function

$$T : \mathbb{R}^k \to V$$

defined above is an isomorphism. In particular:

1. $\dim V = k$.
2. *If $\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_k\}$ is any basis of $\mathbb{R}^k$, then $\{T(\mathbf{v}_1),\ \dots,\ T(\mathbf{v}_k)\}$ is a basis of V.*

Proof. (1) and (2) will follow from Theorem 7.3.1 and Theorem 7.3.2 as soon as we show that T is an isomorphism. Given $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^k$, write $\mathbf{v} = (v_0,\ v_1,\ \dots,\ v_{k-1})$ and $\mathbf{w} = (w_0,\ w_1,\ \dots,\ w_{k-1})$. The first k terms of $T(\mathbf{v})$ and $T(\mathbf{w})$ are $v_0,\ v_1,\ \dots,\ v_{k-1}$ and $w_0,\ w_1,\ \dots,\ w_{k-1}$, respectively, so the first k terms of $T(\mathbf{v}) + T(\mathbf{w})$ are $v_0 + w_0,\ v_1 + w_1,\ \dots,\ v_{k-1} + w_{k-1}$. Because these terms agree with the first k terms of $T(\mathbf{v} + \mathbf{w})$, Lemma 7.5.1 implies that $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$. The proof that $T(r\mathbf{v}) + rT(\mathbf{v})$ is similar, so T is linear.

Now let $[x_n)$ be any sequence in V, and let $\mathbf{v} = (x_0,\ x_1,\ \dots,\ x_{k-1})$. Then the first k terms of $[x_n)$ and $T(\mathbf{v})$ agree, so $T(\mathbf{v}) = [x_n)$. Hence T is onto. Finally, if $T(\mathbf{v}) = [0)$ is the zero sequence, then the first k terms of $T(\mathbf{v})$ are all zero (*all* terms of $T(\mathbf{v})$ are zero!) so $\mathbf{v} = \mathbf{0}$. This means that $\ker T = \{\mathbf{0}\}$, so T is one-to-one. □

Example 7.5.2

Show that the sequences $[1)$, $[n)$, and $[(-1)^n)$ are a basis of the space V of all solutions of the recurrence

$$x_{n+3} = -x_n + x_{n+1} + x_{n+2}$$

Then find the solution satisfying $x_0 = 1$, $x_1 = 2$, $x_2 = 5$.

Solution. The verifications that these sequences satisfy the recurrence (and hence lie in V) are left to the reader. They are a basis because $[1) = T(1,\ 1,\ 1)$, $[n) = T(0,\ 1,\ 2)$, and $[(-1)^n) = T(1,\ -1,\ 1)$; and $\{(1,\ 1,\ 1),\ (0,\ 1,\ 2),\ (1,\ -1,\ 1)\}$ is a basis of $\mathbb{R}^3$. Hence the sequence $[x_n)$ in V satisfying $x_0 = 1$, $x_1 = 2$, $x_2 = 5$ is a linear combination of this basis:

$$[x_n) = t_1[1) + t_2[n) + t_3[(-1)^n)$$

The nth term is $x_n = t_1 + nt_2 + (-1)^n t_3$, so taking $n = 0,\ 1,\ 2$ gives

$$\begin{aligned} 1 = x_0 &= t_1 + \ \ 0 \ \ + t_3 \\ 2 = x_1 &= t_1 + \ t_2 \ - t_3 \\ 5 = x_2 &= t_1 + 2t_2 + t_3 \end{aligned}$$

This has the solution $t_1 = t_3 = \frac{1}{2}$, $t_2 = 2$, so $x_n = \frac{1}{2} + 2n + \frac{1}{2}(-1)^n$.

This technique clearly works for any linear recurrence of length k: Simply take your favourite basis $\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_k\}$ of $\mathbb{R}^k$—perhaps the standard basis—and compute $T(\mathbf{v}_1),\ \dots,\ T(\mathbf{v}_k)$. This is a basis of V all right, but the nth term of $T(\mathbf{v}_i)$ is not usually given as an explicit function of n. (The basis in Example 7.5.2 was carefully chosen so that the nth terms of the three sequences were 1, n, and $(-1)^n$, respectively, each a simple function of n.)

However, it turns out that an explicit basis of V can be given in the general situation. Given the recurrence (7.5) again:

$$x_{n+k} = r_0 x_n + r_1 x_{n+1} + \cdots + r_{k-1} x_{n+k-1}$$

the idea is to look for numbers λ such that the power sequence $[\lambda^n)$ satisfies (7.5). This happens if and only if

$$\lambda^{n+k} = r_0\lambda^n + r_1\lambda^{n+1} + \cdots + r_{k-1}\lambda^{n+k-1}$$

holds for all $n \geq 0$. This is true just when the case $n = 0$ holds; that is,

$$\lambda^k = r_0 + r_1\lambda + \cdots + r_{k-1}\lambda^{k-1}$$

The polynomial

$$p(x) = x^k - r_{k-1}x^{k-1} - \cdots - r_1 x - r_0$$

is called the polynomial **associated** with the linear recurrence (7.5). Thus every root λ of $p(x)$ provides a sequence $[\lambda^n)$ satisfying (7.5). If there are k distinct roots, the power sequences provide a basis. Incidentally, if $\lambda = 0$, the sequence $[\lambda^n)$ is $1,\ 0,\ 0,\ \dots$; that is, we accept the convention that $0^0 = 1$.

Theorem 7.5.2

Let $r_0, r_1, \dots, r_{k-1}$ be real numbers; let

$$V = \{[x_n) \mid x_{n+k} = r_0 x_n + r_1 x_{n+1} + \cdots + r_{k-1} x_{n+k-1} \text{ for all } n \geq 0\}$$

denote the vector space of all sequences satisfying the linear recurrence relation determined by $r_0, r_1, \dots, r_{k-1}$; and let

$$p(x) = x^k - r_{k-1}x^{k-1} - \cdots - r_1 x - r_0$$

denote the polynomial associated with the recurrence relation. Then

1. *$[\lambda^n)$ lies in V if and only if λ is a root of $p(x)$.*
2. *If $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct real roots of $p(x)$, then $\{[\lambda_1^n), [\lambda_2^n), \dots, [\lambda_k^n)\}$ is a basis of V.*

Proof. It remains to prove (2). But $[\lambda_i^n) = T(\mathbf{v}_i)$ where $\mathbf{v}_i = (1, \lambda_i, \lambda_i^2, \dots, \lambda_i^{k-1})$, so (2) follows by Theorem 7.5.1, provided that $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ is a basis of $\mathbb{R}^k$. This is true provided that the matrix with the $\mathbf{v}_i$ as its rows

$$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \cdots & \lambda_k^{k-1} \end{bmatrix}$$

is invertible. But this is a Vandermonde matrix and so is invertible if the λ_i are distinct (Theorem 3.2.7). This proves (2). □

Example 7.5.3

Find the solution of $x_{n+2} = 2x_n + x_{n+1}$ that satisfies $x_0 = a$, $x_1 = b$.

Solution. The associated polynomial is $p(x) = x^2 - x - 2 = (x-2)(x+1)$. The roots are $\lambda_1 = 2$ and $\lambda_2 = -1$, so the sequences $[2^n)$ and $[(-1)^n)$ are a basis for the space of solutions by Theorem 7.5.2. Hence every solution $[x_n)$ is a linear combination

$$[x_n) = t_1[2^n) + t_2[(-1)^n)$$

This means that $x_n = t_1 2^n + t_2(-1)^n$ holds for $n = 0, 1, 2, \dots$, so (taking $n = 0, 1$) $x_0 = a$ and $x_1 = b$ give

$$\begin{aligned} t_1 + t_2 &= a \\ 2t_1 - t_2 &= b \end{aligned}$$

These are easily solved: $t_1 = \frac{1}{3}(a+b)$ and $t_2 = \frac{1}{3}(2a-b)$, so

$$t_n = \tfrac{1}{3}\left[(a+b)2^n + (2a-b)(-1)^n\right]$$

The Shift Operator

If $p(x)$ is the polynomial associated with a linear recurrence relation of length k, and if $p(x)$ has k distinct roots $\lambda_1, \lambda_2, \dots, \lambda_k$, then $p(x)$ factors completely:

$$p(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$$

Each root λ_i provides a sequence $[\lambda_i^n)$ satisfying the recurrence, and they are a basis of V by Theorem 7.5.2. In this case, each λ_i has multiplicity 1 as a root of $p(x)$. In general, a root λ has **multiplicity** m if $p(x) = (x-\lambda)^m q(x)$, where $q(\lambda) \neq 0$. In this case, there are fewer than k distinct roots and so fewer than k sequences $[\lambda^n)$ satisfying the recurrence. However, we can still obtain a basis because, if λ has multiplicity m (and $\lambda \neq 0$), it provides m linearly independent sequences that satisfy the recurrence. To prove this, it is convenient to give another way to describe the space V of all sequences satisfying a given linear recurrence relation.

Let $\mathbf{S}$ denote the vector space of *all* sequences and define a function

$$S : \mathbf{S} \to \mathbf{S} \quad \text{by} \quad S[x_n) = [x_{n+1}) = [x_1,\ x_2,\ x_3,\ \dots)$$

S is clearly a linear transformation and is called the **shift operator** on $\mathbf{S}$. Note that powers of S shift the sequence further: $S^2[x_n) = S[x_{n+1}) = [x_{n+2})$. In general,

$$S^k[x_n) = [x_{n+k}) = [x_k,\ x_{k+1},\ \dots) \quad \text{for all } k = 0,\ 1,\ 2,\ \dots$$

But then a linear recurrence relation

$$x_{n+k} = r_0x_n + r_1x_{n+1} + \cdots + r_{k-1}x_{n+k-1} \quad \text{for all } n = 0,\ 1,\ \dots$$

can be written

$$S^k[x_n) = r_0[x_n) + r_1S[x_n) + \cdots + r_{k-1}S^{k-1}[x_n) \tag{7.6}$$

Now let $p(x) = x^k - r_{k-1}x^{k-1} - \cdots - r_1x - r_0$ denote the polynomial associated with the recurrence relation. The set $\mathbf{L}[\mathbf{S},\ \mathbf{S}]$ of all linear transformations from $\mathbf{S}$ to itself is a vector space (verify[8]) that is closed under composition. In particular,

$$p(S) = S^k - r_{k-1}S^{k-1} - \cdots - r_1S - r_0$$

is a linear transformation called the **evaluation** of p at S. The point is that condition (7.6) can be written as

$$p(S)\{[x_n)\} = 0$$

In other words, the space V of all sequences satisfying the recurrence relation is just $\ker[p(S)]$. This is the first assertion in the following theorem.

Theorem 7.5.3

Let $r_0,\ r_1,\ \dots,\ r_{k-1}$ be real numbers, and let

$$V = \{[x_n) \mid x_{n+k} = r_0x_n + r_1x_{n+1} + \cdots + r_{k-1}x_{n+k-1} \quad \text{for all } n \geq 0\}$$

[8] See Exercises 9.1.19 and 9.1.20.

denote the space of all sequences satisfying the linear recurrence relation determined by $r_0, r_1, \ldots, r_{k-1}$. *Let*

$$p(x) = x^k - r_{k-1}x^{k-1} - \cdots - r_1 x - r_0$$

denote the corresponding polynomial. Then:

1. $V = \ker[p(S)]$, *where* S *is the shift operator.*

2. *If* $p(x) = (x-\lambda)^m q(x)$, *where* $\lambda \neq 0$ *and* $m > 1$, *then the sequences*

$$\{[\lambda^n), [n\lambda^n), [n^2\lambda^n), \ldots, [n^{m-1}\lambda^n)\}$$

all lie in V *and are linearly independent.*

Proof (Sketch). It remains to prove (2). If $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$ denotes the binomial coefficient, the idea is to use (1) to show that the sequence $s_k = \left[\binom{n}{k}\lambda^n\right)$ is a solution for each $k = 0, 1, \ldots, m-1$. Then (2) of Theorem 7.5.1 can be applied to show that $\{s_0, s_1, \ldots, s_{m-1}\}$ is linearly independent. Finally, the sequences $t_k = [n^k\lambda^n)$, $k = 0, 1, \ldots, m-1$, in the present theorem can be given by $t_k = \sum_{j=0}^{m-1} a_{kj}s_j$, where $A = \left[a_{ij}\right]$ is an invertible matrix. Then (2) follows. We omit the details. □

This theorem combines with Theorem 7.5.2 to give a basis for V when $p(x)$ has k real roots (not necessarily distinct) none of which is zero. This last requirement means $r_0 \neq 0$, a condition that is unimportant in practice (see Remark 1 below).

Theorem 7.5.4

Let $r_0, r_1, \ldots, r_{k-1}$ *be real numbers with* $r_0 \neq 0$; *let*

$$V = \{[x_n) \mid x_{n+k} = r_0 x_n + r_1 x_{n+1} + \cdots + r_{k-1}x_{n+k-1} \text{ for all } n \geq 0\}$$

denote the space of all sequences satisfying the linear recurrence relation of length k *determined by* $r_0, \ldots, r_{k-1}$; *and assume that the polynomial*

$$p(x) = x^k - r_{k-1}x^{k-1} - \cdots - r_1 x - r_0$$

factors completely as

$$p(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_p)^{m_p}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ *are distinct real numbers and each* $m_i \geq 1$. *Then* $\lambda_i \neq 0$ *for each* i, *and*

$$\begin{array}{c}
[\lambda_1^n), [n\lambda_1^n), \ldots, [n^{m_1-1}\lambda_1^n) \\
[\lambda_2^n), [n\lambda_2^n), \ldots, [n^{m_2-1}\lambda_2^n) \\
\vdots \\
[\lambda_p^n), [n\lambda_p^n), \ldots, [n^{m_p-1}\lambda_p^n)
\end{array}$$

is a basis of V.

Proof. There are $m_1 + m_2 + \cdots + m_p = k$ sequences in all so, because $\dim V = k$, it suffices to show that they are linearly independent. The assumption that $r_0 \neq 0$, implies that 0 is not a root of $p(x)$. Hence each $\lambda_i \neq 0$, so $\{[\lambda_i^n), [n\lambda_i^n), \dots, [n^{m_i-1}\lambda_i^n)\}$ is linearly independent by Theorem 7.5.3. The proof that the whole set of sequences is linearly independent is omitted. □

Example 7.5.4

Find a basis for the space V of all sequences $[x_n)$ satisfying

$$x_{n+3} = -9x_n - 3x_{n+1} + 5x_{n+2}$$

Solution. The associated polynomial is

$$p(x) = x^3 - 5x^2 + 3x + 9 = (x-3)^2(x+1)$$

Hence 3 is a double root, so $[3_n)$ and $[n3^n)$ both lie in V by Theorem 7.5.3 (the reader should verify this). Similarly, $\lambda = -1$ is a root of multiplicity 1, so $[(-1)^n)$ lies in V. Hence $\{[3^n), [n3^n), [(-1)^n)\}$ is a basis by Theorem 7.5.4.

Remark 1

If $r_0 = 0$ [so $p(x)$ has 0 as a root], the recurrence reduces to one of shorter length. For example, consider

$$x_{n+4} = 0x_n + 0x_{n+1} + 3x_{n+2} + 2x_{n+3} \tag{7.7}$$

If we set $y_n = x_{n+2}$, this recurrence becomes $y_{n+2} = 3y_n + 2y_{n+1}$, which has solutions $[3^n)$ and $[(-1)^n)$. These give the following solution to (7.5):

$$[0, 0, 1, 3, 3^2, \dots)$$
$$[0, 0, 1, -1, (-1)^2, \dots)$$

In addition, it is easy to verify that

$$[1, 0, 0, 0, 0, \dots)$$
$$[0, 1, 0, 0, 0, \dots)$$

are also solutions to (7.7). The space of all solutions of (7.5) has dimension 4 (Theorem 7.5.1), so these sequences are a basis. This technique works whenever $r_0 = 0$.

Remark 2

Theorem 7.5.4 completely describes the space V of sequences that satisfy a linear recurrence relation for which the associated polynomial $p(x)$ has all real roots. However, in many cases of interest, $p(x)$ has complex roots that are not real. If $p(\mu) = 0$, μ complex, then $p(\overline{\mu}) = 0$ too ($\overline{\mu}$ the conjugate), and the main observation is that $[\mu^n + \overline{\mu}^n)$ and $[i(\mu^n + \overline{\mu}^n))$ are *real* solutions. Analogs of the preceding theorems can then be proved.

Exercises for 7.5

Exercise 7.5.1 Find a basis for the space V of sequences $[x_n)$ satisfying the following recurrences, and use it to find the sequence satisfying $x_0 = 1$, $x_1 = 2$, $x_2 = 1$.

a. $x_{n+3} = -2x_n + x_{n+1} + 2x_{n+2}$

b. $x_{n+3} = -6x_n + 7x_{n+1}$

c. $x_{n+3} = -36x_n + 7x_{n+2}$

Exercise 7.5.2 In each case, find a basis for the space V of all sequences $[x_n)$ satisfying the recurrence, and use it to find x_n if $x_0 = 1$, $x_1 = -1$, and $x_2 = 1$.

a. $x_{n+3} = x_n + x_{n+1} - x_{n+2}$

b. $x_{n+3} = -2x_n + 3x_{n+1}$

c. $x_{n+3} = -4x_n + 3x_{n+2}$

d. $x_{n+3} = x_n - 3x_{n+1} + 3x_{n+2}$

e. $x_{n+3} = 8x_n - 12x_{n+1} + 6x_{n+2}$

Exercise 7.5.3 Find a basis for the space V of sequences $[x_n)$ satisfying each of the following recurrences.

a. $x_{n+2} = -a^2x_n + 2ax_{n+1}$, $a \neq 0$

b. $x_{n+2} = -abx_n + (a+b)x_{n+1}$, $(a \neq b)$

Exercise 7.5.4 In each case, find a basis of V.

a. $V = \{[x_n) \mid x_{n+4} = 2x_{n+2} - x_{n+3}$, for $n \geq 0\}$

b. $V = \{[x_n) \mid x_{n+4} = -x_{n+2} + 2x_{n+3}$, for $n \geq 0\}$

Exercise 7.5.5 Suppose that $[x_n)$ satisfies a linear recurrence relation of length k. If $\{\mathbf{e}_0 = (1, 0, \dots, 0)$, $\mathbf{e}_1 = (0, 1, \dots, 0), \dots, \mathbf{e}_{k-1} = (0, 0, \dots, 1)\}$ is the standard basis of $\mathbb{R}^k$, show that

$$x_n = x_0T(\mathbf{e}_0) + x_1T(\mathbf{e}_1) + \cdots + x_{k-1}T(\mathbf{e}_{k-1})$$

holds for all $n \geq k$. (Here T is as in Theorem 7.5.1.)

Exercise 7.5.6 Show that the shift operator S is onto but not one-to-one. Find $\ker S$.

Exercise 7.5.7 Find a basis for the space V of all sequences $[x_n)$ satisfying $x_{n+2} = -x_n$.